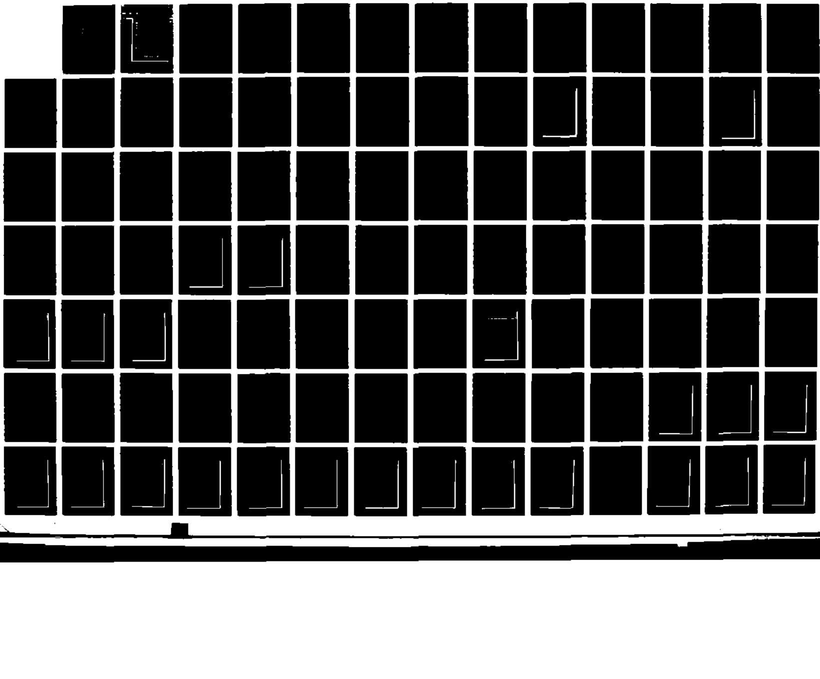
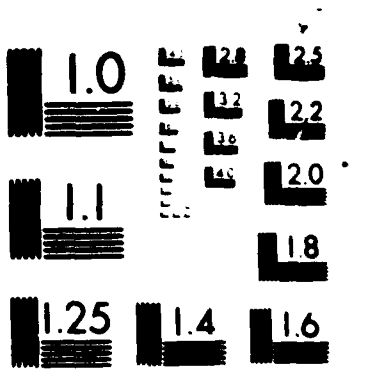
AD-A183 565 ORGANIZATIONAL PRODUCTIVITY MEASUREMENT: THE DEVELOPMENT AND EVALUATION O..(U) HOUSTON UNIV TX DEPT OF PSYCHOLOGY R D PRITCHARD ET AL. JUL 87 1/2

UNCLASSIFIED AFHRL-TR-86-64 F41689-83-C-0039 F/G 5/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS

DTIC FILE COPY

AFHRL-TR-86-64

(12)

AIR FORCE

HUMAN RESOURCES

AD-A183 565

ORGANIZATIONAL PRODUCTIVITY MEASUREMENT: THE DEVELOPMENT AND EVALUATION OF AN INTEGRATED APPROACH

Robert D. Pritchard
Steven D. Jones
Philip L. Roth
Karla K. Stuebing
Steven E. Ekeberg

Institute for Organizational Behavior Research
and
Department of Psychology
University of Houston
Houston, Texas 77004

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

July 1987
Interim Report for Period January 1984 - July 1986

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

ADA183565

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION: Unclassified

1b RESTRICTIVE MARKINGS:

2a SECURITY CLASSIFICATION AUTHORITY:

2b DECLASSIFICATION / DOWNGRADING SCHEDULE:

3 DISTRIBUTION / AVAILABILITY OF REPORT: Approved for public release; distribution is unlimited.

4 PERFORMING ORGANIZATION REPORT NUMBER(S):

5 MONITORING ORGANIZATION REPORT NUMBER(S): AFHRL-TR-86-64

6a NAME OF PERFORMING ORGANIZATION: Department of Psychology

6b OFFICE SYMBOL *(If applicable)*:

7a NAME OF MONITORING ORGANIZATION: Manpower and Personnel Division

6c ADDRESS *(City, State, and ZIP Code)*:
University of Houston
Houston, Texas 77004

7b ADDRESS *(City, State, and ZIP Code)*:
Air Force Human Resources Laboratory
Brooks Air Force Base, Texas 78235-5601

8a NAME OF FUNDING / SPONSORING ORGANIZATION: Air Force Human Resources Laboratory

8b OFFICE SYMBOL *(If applicable)*: HQ AFHRL

9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F41689-83-C-0039

8c ADDRESS *(City, State, and ZIP Code)*:
Brooks Air Force Base, Texas 78235-5601

10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
62703F	7719	23	01

11 TITLE *(Include Security Classification)*:
Organizational Productivity Measurement: The Development and Evaluation of an Integrated Approach

12 PERSONAL AUTHOR(S):
Pritchard, R.D.; Jones, S.D.; Roth, P.L.; Stuebing, K.K.; Ekeberg, S.E.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Interim	FROM Jan 84 TO Jul 86	July 1987	136

16 SUPPLEMENTARY NOTATION:

17 COSATI CODES

FIELD	GROUP	SUB-GROUP
05	03	
05	08	

18 SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*:
effectiveness
effectiveness measurement
efficiency
efficiency measurement
productivity
productivity measurement

19 ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This report describes a field evaluation of a new approach to the measurement and enhancement of organizational productivity. In order to enhance productivity in the Air Force, it is necessary to have good productivity measures. A newly developed measurement methodology is presented and its application to Air Force organizations is evaluated. The system was developed for five maintenance and supply sections at an operational Air Force base. After helping develop the measurement system, each unit received productivity feedback from it for 5 months. Next, goal setting was added and feedback provided for another 5 months. Finally, incentives were added to feedback and goal setting. Once the productivity measurement system was developed, it became the basis for providing formal monthly feedback about unit productivity to personnel and their managers. The effectiveness of the system was evaluated during the development and implementation periods. Results showed the approach to be a very effective method of productivity measurement and enhancement. The system was implemented effectively, unit personnel were cooperative in developing and using it, and it showed good psychometric characteristics. Compared to the baseline period, the feedback that was produced by the system resulted in an average productivity gain over baseline of 50% during feedback, 75% during feedback plus goal setting, and 76% when incentives were

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT:
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION:

22a NAME OF RESPONSIBLE INDIVIDUAL: Nancy J. Allin, Chief, STINFO Office

22b TELEPHONE *(Include Area Code)*: (512) 536-3877

22c OFFICE SYMBOL: AFHRL/TSR

DD Form 1473, JUN 86 — *Previous editions are obsolete* — SECURITY CLASSIFICATION OF THIS PAGE

Unclassified

19. (Concluded)

added. The positive effects lasted over time, and continued after the departure of the research team. The incumbents and supervisors of the units evaluated the system very positively. At the end of the research, the units continued to use the system on their own, and management has requested that it be used in other units at the base.

The researchers concluded that the system worked well because it provided a single index of productivity along with subindices, was well accepted by unit personnel, was a valid measure of productivity, allowed for direct comparison across units, provided for greater unit accountability, allowed unit personnel to focus on common objectives, and, most importantly, increased motivation. In addition, the system allowed for the tracking of productivity over time so that the effects of organizational changes could be assessed. It is flexible in that it can be used with both effectiveness and efficiency approaches to organizational productivity. It can accommodate changes in organizational practices, policies, and priorities over time. It can be applied to any level of organization, and allows different units to be combined into one measurement system and be directly compared. It also affords a way of measuring the effectiveness of management, and shows promise for use in much larger organizational units. The basic measurement and aggregation strategy also has applications in management information systems, performance appraisal, and other situations where multiple sources of data must be combined into an overall index or judgment.

SUMMARY

This technical paper describes the results of a field evaluation of a new approach to the measurement of organizational productivity. This approach involves (a) identifying the objectives of the unit, (b) identifying measures or indicators of how well the unit is meeting these objectives, and (c) developing functional relationships between performance on the indicators and the contribution that those levels of the indicators make to overall effectiveness.

The productivity measurement system was developed for five sections in maintenance and supply at an operational Air Force base. The productivity measures derived from the system were used as a basis for monthly feedback to the units for a period of five months. After this period, goal setting was added to the feedback for five months. Finally, incentives were added to the feedback and goal setting.

Results showed the system to be a very effective method of productivity measurement and enhancement. Its implementation was effective, unit personnel were cooperative in developing and using it, and it showed good psychometric characteristics. Using the feedback that was produced by the system resulted in an average gain in productivity of 50% over baseline, across the five units. When goal setting was added, the mean increase was 75% over baseline. When incentives were added, the mean increase was 76% over baseline. The positive effects lasted over time, and continued after the departure of the research team. The incumbents and supervisors of the units evaluated the system very positively. After the research was completed, the units continued to use the system on their own, and managers have requested that it be used in other units at the base.

This approach shows promise for use in much larger organizational units. The basic measurement and aggregation strategy also has applications in management information systems, criterion development, test validation, measures of managerial performance, performance appraisal, and other situations where multiple sources of data must be combined into an overall index or judgment.

PREFACE

We wish to thank the following people on our staff who assisted us with their many contributions to this effort: Judy Moore, Steve Schweigert, Molly Jackson, Todd Lambertus, Anne Davee, Mark Tubbs, Frances Svyantek, Donna Payne, Janet Hennessy, and Patricia Galgay. In addition, we would like to thank all Air Force personnel who were involved in the project. We shall not name them individually, since over 100 people participated. However, the efforts of the management and personnel in the Communications-Navigation section and in the Materiel Storage and Distribution branch were integral to the improvements in productivity described in this paper. We also acknowledge the managers above these two organizations who gave us approval to conduct the research in their areas and who also met with us to review the productivity measurement system and the results of the interventions. We also wish to thank our contract monitor, Dr. Charles N. Weaver of the Air Force Human Resources Laboratory, for his support of this project both in terms of his ideas and his advocacy.

TABLE OF CONTENTS

List of Figures

List of Tables

I. INTRODUCTION

The problem of enhancing productivity has been a major concern for some time. In the popular media and in technical presentations, enhancing productivity has been seen as an issue that has implications for our quality of life, our economy, and our competitive position in the world marketplace (Alluisi & Meigs, 1983; American Productivity Center, 1981). In addition, individual organizations are continually concerned about increasing their productivity in order to improve their operational effectiveness. This concern for increasing productivity is shared by the Air Force, and has led the Air Force Human Resources Laboratory (AFHRL) to explore ways of enhancing productivity.

As important as increasing productivity is, attempts to do so are usually hampered by the lack of good measures of productivity. The basic idea is a simple one. To increase productivity, we must first be able to measure it.

The purpose of this technical report is to describe a new approach to measuring organizational productivity that we feel is a substantial improvement over existing methods, and to present the results of an implementation of this system in an operational Air Force environment. This report is one of three reports coming from the project. One of the other reports (Pritchard, Jones, Roth, Stuebing, and Ekeberg, 1987) is a research report focusing on the results of the feedback, goal setting, and incentive interventions. The third report (Pritchard, Stuebing, Jones, Roth, and Ekeberg, 1987) is not a research report, but a manager's manual with practical directions for designing and implementing feedback, goal setting, and incentive systems.

In the remainder of this report we shall (a) briefly review approaches to organizational productivity, (b) present our approach to organizational productivity, (c) discuss the results of a field test that utilized this approach, (d) present the positive features of this approach and (e) draw conclusions about the effort.

II. APPROACHES TO ORGANIZATIONAL PRODUCTIVITY

Although much has been written on the subject of organizational productivity, there is little consensus concerning its definition (Tuttle, 1983). Such a lack of consensus is perhaps not surprising since there are many approaches to and perspectives on productivity. There are, however, several major issues that should be addressed. In the review below, we will first present these issues and then briefly discuss specific techniques that have been used to measure productivity.

Efficiency vs. Effectiveness

The first issue to be addressed is whether an efficiency or an effectiveness approach should be used in measuring productivity. Both have been proposed and used. Efficiency is typically thought of as an output-to-input ratio. For example, monthly manufacturing output divided by manpower used to produce that output would be an efficiency measure. Effectiveness is usually discussed as the relationship of outputs to some standard or expectation. For example, monthly manufacturing output expressed as a percentage of the goal for that month would be an effectiveness measure. In addition, effectiveness usually includes quality of the output as well as quantity.

Efficiency is the more widely used of the two concepts because it is easier to measure and standardize across organizations, industries, and nations (Norman & Bahiri, 1972). When we hear that productivity growth in the United States has declined over the last 20 years (American Productivity Center, 1981), it is an efficiency ratio that is being quoted (i.e., price deflated gross national product divided by worker hours). Effectiveness is a much broader concept because it includes other concepts such as standards, objectives of the organization, expectations of interested parties (e.g., shareholders, regulatory agencies, and customers), and the viability of the organization relative to its competition. As Mendelow (1983) argued: "The most efficient slide-rule manufacturer would be out of business today assuming it had not adapted its product line to meet the onslaught of hand-held calculators " (p. 70). Proponents of the effectiveness concept argue that as complexity and ambiguity of the work increase, effectiveness measures become more important than efficiency (Balk, 1975). Effectiveness can also reflect the organization's bargaining position relative to its environment, whereas the efficiency concept does not.

Some authors define productivity as a combination of efficiency and effectiveness (Balk, 1975; Coulter, 1979; Hanes & Kriebel, 1978; National Center for Productivity and Quality of Working Life, 1978; Sibson, 1976; Tuttle, 1981). Typical formulations (Balk, 1975) are:

Productivity = Efficiency + Effectiveness or

Productivity = Output/Input + Output/Standard

Perspective Taken

A second issue in the productivity literature is that although differing approaches to productivity can be understood in terms of an efficiency, effectiveness, or an integrated scheme, they can also be understood in terms of the perspective taken. That is, the approach to measuring productivity is determined by the perspective of those doing the measuring. Tuttle (1983) suggested five perspectives:

1. **Economist's Perspective.** In this approach, productivity is output divided by associated inputs such as labor, capital, intermediate products purchased, and time. This approach is typically applied to very macro units such as whole industries or countries.

2. **Engineer's Perspective.** In this approach, productivity is equated with the efficiency of the operation, based on a comparison of energy as the input and useful work as the output. This approach would typically be used with one organization, or a part of the organization. In addition, it would typically focus on the equipment/hardware aspects of the organization.

3. **Accountant's Perspective.** Here the focus is on the financial performance of the organization. Various ratios such as profit divided by sales would be examined.

4. **Manager's Perspective.** This is a broad definition of productivity which includes quality, quantity, disruption, turnover, and absenteeism (Katzell, Yankelovich, Fein, Ornati, & Nash, 1975).

5. **Industrial/Organizational Psychologist's Perspective.** Here the concern is primarily the personnel subsystem of the organization, and the efficiency or effectiveness of that subsystem.

Clearly, these approaches are quite different. They measure different things, and they are used for different purposes. They would also result in very different productivity measurement systems.

<u>Organizational Model Used</u>

A third issue is that the measurement of productivity will be determined by the organizational model used. This issue comes from the literature on organizational effectiveness. Campbell (1977) summarized the theoretical viewpoints on organizational effectiveness. While there are numerous models of organizational effectiveness (Campbell, Bownas, Peterson, & Dunnette, 1974; Coulter, 1979; Engel, 1977; Goodman & Pennings, 1976; Mahoney & Frost, 1974; Mahoney &

Weitzel, 1969; Price, 1968; Steers, Porter, Mowday, & Stone, 1975), Campbell (1977) identified two general models:

1. **Goal Centered.** This model assumes that the way to assess organizational effectiveness is to develop criterion measures assessing how well the organization's goals are being achieved. The goals to be assessed are referred to as operative goals. These goals are the ends sought through the actual operating policies of the organization (Perrow, 1961), as opposed to the officially stated goals of the organization (Keeley, 1978).

2. **Natural Systems.** This model assumes that the demands from the environment are so dynamic and complex that a finite number of goals cannot be defined. The theory holds that the organization should have the overall goal of maintaining its viability, without depleting its environment. This view focuses on the means of obtaining organizational viability such as internal consistency, judicious distribution and use of resources, etc. The focus is on the people in the organization, not on the state of the organization's technology or its physical structure.

These two global perspectives of organizations are quite different from one another, and imply quite different approaches to measuring organizational productivity.

<u>What To Include In The Measurement</u>

The next issue is what measures should be included in the measure of productivity. Clearly, the different perspectives such as the economist's and the accountant's have implications for what measures are included, as does the organizational model used. There are, however, a variety of other possibilities. Campbell (1977), for example, listed 30 types of measures that have been used.

1. Overall Effectiveness
2. Productivity
3. Efficiency
4. Profit
5. Quality
6. Accidents
7. Growth
8. Absenteeism
9. Turnover

10. Job Satisfaction
11. Motivation
12. Morale
13. Control
14. Conflict/Cohesion
15. Flexibility/Adaptation
16. Planning and Goal Setting
17. Goal Consensus
18. Internalization of Organizational Goals
19. Role and Norm Congruence
20. Managerial Interpersonal Skills
21. Managerial Task Skills
22. Information Management and Communications
23. Readiness
24. Utilization of Environment
25. Evaluations by External Entities
26. Stability
27. Value of Human Resources
28. Participation and Shared Influence
29. Training and Development Emphasis
30. Achievement Emphasis

Seashore and Yuchtman (1967) reported on a factor analysis of organizational productivity scores for insurance agencies. They identified ten factors, many of which were quite different from those listed by Campbell. They included new member productivity, youthfulness of members, business mix, manpower growth, and market penetration.

The variety of measures that could be included in a productivity measurement system clearly shows that no one set of measures constitutes productivity. The diversity of possible measures must be considered in the design of a productivity measurement system.

Completeness Of The Measurement System

In contrast to the issue of what constitutes productivity, there is considerable agreement that a productivity measurement system should include all important aspects of the organization's work. If the system is not complete, it could easily encourage neglect of those organizational objectives that are not included as part of the measurement system. In such a situation, the overall effectiveness of the organization would suffer.

Duerr (1974) reported several instances of this problem. One case involved a manufacturing company. The plants in the company were compared, and incentives were given to supervisors and managers on the basis of an operating efficiency ratio. Certain aspects of productivity that were not included in this ratio were allowed to suffer so that the operating efficiency ratios would be favorable. The result was that the company sustained significant economic loss because certain measures were neglected.

One Air Force training manager described the effects of an incomplete productivity measurement system this way:

> We measure the things that are easy to measure but the greatest part of a mission is constituted by things that are not easy to measure ... [we] measure those things that are easy to measure -- and we measure them just great -- but whether we achieve our mission will depend on these other objectives that we've avoided because we're not sure how to measure them. Until we come to grips with this sort of thing, we're not sure we're assessing the right thing. And if you don't assess the right thing you'll never know if you're productive (Tuttle, 1981, p. 24-25).

Use Of An Overall Index Of Productivity

Another broad issue for productivity measurement systems is the use of an overall index of productivity. We would argue that the use of a single index is very important because of its motivational value. A single index provides the members of the unit with a sense of improvement or decrement. The single index would also seem beneficial for information purposes. A large number of pieces of information on organizational functioning can be very difficult to assimilate and use for making decisions.

The problem of integrating numerous measures of productivity into a composite is similar to the problem of developing a composite measure of individual performance when several dimensions of performance are identified. Schmidt and Kaplan (1971) reviewed the controversy between composite and multiple criteria. Several points were made, two of which are relevant for productivity measurement:

1. Composite criteria are useful for decision making, providing the weights applied to each of the criteria contained in the composite are subject to a verification procedure.

2. The combining of unrelated variables into a composite presents a problem if the variables do not represent an underlying dimension such as an economic value or effectiveness. Apples and oranges can be added only if they can logically be converted to a measure of their underlying dimension.

There is, however, an additional problem with the weighting of individual measures which concerns the linearity of the weighting. Giving a weight to a productivity measure according to its importance assumes that its relative importance is the same regardless of the level of performance on that measure. The fact that this may not be a valid assumption has been pointed out by Campbell (1977):

> ...in the real world it is probably a mistake to think of effectiveness criterion variables, regardless of how many there are or at what level they are, in terms of continuous and linear functions. For example, higher and higher retention rates may be "good" up to a point and then become "bad " (p. 44).

These points suggest that combining or translating individual indices into a composite measure of overall productivity is of considerable value because of the motivational and informational advantages. It is also a reasonable goal since the individual measures form a clear underlying dimension: the productivity of the organization. It is, however, important to accomplish this in such a way that accurate relative weights of the individual measures are maintained, and non-linearities are preserved.

Examples Of Productivity Measurement Systems

Several authors have presented methods for measuring productivity. These cut across many of the conceptual issues that have been presented above. Five of these methods will be presented here.

One method is called Total Performance Measurement (Joint Financial Management Improvement Program, 1976), and is a technique for measuring productivity that has been used in a number of Government organizations. This technique combines industrial engineering and behavioral science technologies to measure various aspects of productivity. Objective productivity indices, such as efficiency ratios, are collected. Questionnaires are designed and administered to customers of the organization as well as its employees. The questionnaire data are related to the "hard" productivity indices, and these data are then presented to management in a feedback session. The feedback session is designed to reveal the causes of good and poor organizational performance, and to form action plans.

A second method was presented by Peeples (1978), who measured productivity in 14 data processing centers. In this system, management weighted each aspect of productivity by assigning points for different levels of goal attainment. The more important goals could earn more total points than less important goals. Within each goal, more points could be earned as the organization approached the goal attainment level. The total number of points was the composite measure of productivity for each data processing center. The centers competed against one another for recognition in terms of total points earned. Significant increases in productivity and financial indices were reported.

A third approach, by Felix and Riggs (1983), presented a productivity measurement system that depends on an "objectives matrix" to combine all productivity criteria into a composite index. The objectives matrix matches different levels of performance on each productivity criterion with a performance level that ranges from one to ten. The highest goal level reasonably possible for each criterion would receive a performance score of 10. The lowest likely score for each criterion would receive a performance score of 0, and likewise for the points in between. Each criterion is weighted for importance, and the performance scores for each criterion are multiplied by the weights. The sum of the weighted performance scores is an overall index of organizational productivity for that unit. For an organization consisting of several units, the authors suggested weighting the overall productivity score for each unit by the number of people employed in each unit. The sum of these weighted overall productivity scores for the units is the aggregated productivity for the larger organization.

A fourth method (Kim, 1980) is a technique for combining effectiveness and efficiency measures. This approach measures effectiveness by dividing each criterion score by the goal level for that criterion. Thus, each criterion has a percent effectiveness value. These percentages for each of the criteria are multiplied by weights reflecting their importance; then they are summed to get a weighted composite effectiveness score. The weighted composite effectiveness score is divided by the number of criteria to get a weighted average composite effectiveness index. This final index is then divided by cost ratios to get effectiveness-to-efficiency ratios.

Finally, Tuttle and his associates (Tuttle, 1981; Tuttle & Weaver, 1986; Tuttle, Wilkinson, & Matthews, 1985) presented a detailed methodology for a participatively developed productivity measurement system. Their approach, known as the Methodology for Generating Efficiency and Effectiveness Measures (MGEEM), begins with the following steps:

1. Management makes the decision to measure productivity.
2. A measurement coordinator from within the organization is selected.
3. Researchers familiarize themselves with the organization.
4. The boundaries of the organization are defined.
5. An organizational diagram is constructed.

The heart of this system requires that managers and employees meet to identify the "key result areas" of the organization's performance, which correspond to the organization's objectives and the support activities for those objectives. Next, organizational members are asked for "indicators" or measures of those key result areas. Tuttle employs the nominal group technique to elicit ideas from supervisors and employees. Thus, the system reflects the ideas of those who will be using the system. Although he does not focus on it in detail, Tuttle suggests that the indicators can be combined into a composite index using the matrix format described by Felix and Riggs (1983).

Conclusions From The Literature

In summary, there are many approaches, perspectives, and issues relevant to productivity measurement. It is tempting to ask, "What is the best definition and perspective to use in conceptualizing productivity?" However, we believe that this is the wrong question. Efficiency and effectiveness approaches both have their place, as do the different perspectives. How one resolves some of the other issues, such as how and what to measure, depends on the circumstances. The better question is, "Under which circumstances is which approach most appropriate?" For different purposes, very different approaches would be used.

III. OUR APPROACH TO ORGANIZATIONAL PRODUCTIVITY

Following this line of reasoning, to establish our approach to organizational productivity, we first need to describe our purpose in measuring it. In the simplest terms, our purpose in measuring productivity is to be able to increase it within a given organization or part of that organization. Our assumption is that the people in the organization have a great impact on the productivity of the organization; that is, what they do and how they do it are most important. Although the technical subsystem is also important, our focus is not on that part of the system directly, but, rather on how the technical subsystem is used by the people.

Therefore, to increase productivity, we need to increase the productivity of the people in the organization.

The mechanism by which this increase would occur is primarily a motivational one. That is, if motivated to do so, personnel would exert more effort and be more persistent in their efforts. They would work more efficiently in the sense that their efforts would be more directly related to organizational objectives. They would improve their work strategies and would use their own and others' time and efforts with less waste. This suggests that the perspective we will be taking is a combination of the manager's perspective and the industrial/organizational psychologist's perspective.

Secondly, although we believe that both efficiency and effectiveness approaches should be included in a productivity measurement system, we believe that the appropriate approach is first to consider productivity as effectiveness rather than efficiency. We take this position for three reasons. First, effectiveness--with its orientation toward goal attainment--is a broader definition of productivity, in that it results in a measurement system that expresses productivity in terms of how good that productivity is. By contrast, an efficiency approach does not carry with it information as to what constitutes a good or bad level of efficiency. The second reason for our adopting the effectiveness approach is that by taking this approach, we can more easily generate a measurement system that can combine all aspects of the organization's productivity into a single measure. The final reason is that the system we are proposing makes it possible to obtain an effectiveness measure and weight it by inputs to arrive at a system that combines the best aspects of both the effectiveness and the efficiency approaches.

The organizational model we will be using is patterned more after the goal centered model than the natural systems model. The natural systems model emphasizes the interaction of the organization with its outside environment. This is clearly an important aspect of the long-range viability of an organization. However, issues such as how the organization will interact with the environment, what environmental changes will be forthcoming, and how to prepare for these changes are matters that are the responsibility of top management. Personnel from middle management down to incumbents do not typically get involved in such broad issues. Top management must make these decisions and then formulate plans and objectives for the rest of the organization to follow so that the organization can successfully interact with its environment.

Therefore, although it may be appropriate to consider the organization based on the natural systems model at the level of top management, the goal centered model is more appropriate if one is focused on the lower levels of the organization. This suggests that if the purpose of the productivity measurement is to increase the motivation of the members of the organization, especially from middle management down, the appropriate model to use is the goal centered model.

Given our purpose of increasing productivity, it is critical that the measurement system be complete, so that increases in measured aspects of the work are not made at the expense of equally important but unmeasured aspects. Finally, the individual measures should be combined into an overall measure of productivity for both motivational and informational purposes. This must be done in a manner that preserves the relative importance of the measures, and captures any non-linearity of the measures.

Description Of The Productivity Measurement System

The theoretical background for this approach to the measurement of organizational productivity stems from the theory of organizational behavior presented by Naylor, Pritchard, and Ilgen (1980). In this theory, an individual's role is seen as a series of relationships, called *contingencies*. These contingencies not only indicate what the important things are that the person must do in the job (called products), but also show the *relationship between* the amount of each of these activities and how that level of the product is evaluated.

This approach to roles has the advantage of indicating more than the typical information present in role specification. The typical information is limited to a listing of the important duties a person must perform on the job. In the Naylor, et al. approach this information is supplemented by the level of performance that is expected in each area, and how positively or negatively each level of performance is evaluated.

In essence, we used the Naylor, et al. conceptual approach of using products and contingencies, and extended its application from individuals to organizational units. This application led to the development of a number of unique features for a productivity measurement system. We shall discuss these later in this paper.

A second source for the development of our approach was based on the work of Tuttle (1981, Tuttle, et al., 1985). In this work, also supported by the Air Force Human Resources Laboratory, Tuttle developed an approach to measuring productivity that included methods of going from what we call products to obtaining objective indicators of how well these products were being produced. He used a variety of group techniques, some of which we used also.

Steps in the Development of the Productivity Measurement System.

The technique to generate the productivity measurement system consists of four distinct steps: (a) identify salient products, (b) develop indicators of these products, (c) establish contingencies, and (d) put the system together.

Step 1: Identify Products.

Every organization has a set of activities that it is expected to perform. These activities result in a set of what Naylor, et al. (1980) called *products*. In using the term "product," we mean more than merely a tangible thing that is produced. Products can be thought of as the set of objectives that the organization is expected to accomplish. The productivity of the organization is a function of how effectively the organization generates these products. The first step in developing the productivity measurement system is to identify these products.

To present the steps involved in developing the productivity measurement system, we shall use an extended hypothetical example that will make each step more concrete. For this example, we shall use a maintenance organization that diagnoses and repairs aircraft electronic communications equipment. The organization's primary responsibility is to repair, as quickly and as accurately as possible, the various items that are brought in when they malfunction. If a repaired item does not function properly when installed in the aircraft, it is returned to them for reaccomplishment of the repair. The unit is periodically inspected by a Quality Control function, which determines whether maintenance personnel are accurately following the procedures for repair detailed in available repair manuals. The maintenance unit also has responsibility for conducting on-the-job training, and a technician can repair a piece of equipment only if he/she has passed the training certification required for that piece of equipment. Thus, it is important that a sufficient number of people be qualified through training so that all the items can be repaired in a timely manner.

To develop the system, the first step would be to meet with people from the organization to identify the salient products. Let us assume that the following products are identified:

1. Quality of repair.

2. Ability to meet demand for repairing items (i.e., the organization's ability to repair the needed equipment quickly).

3. Ability to meet training needs (i.e., the degree to which the organization meets its on-the-job training needs).

In actual fact, there might well be more products in such an organization. However, since our intent here is to explain the logic of the system, we shall use only these three so that the example remains simple enough for clear presentation.

Step 2: Develop Indicators.

Once the products are determined, the next step is to develop *indicators* for each of these products. An indicator is a measure of how well the organization is generating the product in question. The indicators are determined from interaction with the people in the organization, who are asked to think of those things which would show how well people in the organization are producing their products. There may be only one indicator for a given product, or there may be more than one. Some indicators will already be available; some will have to be newly developed. After the indicators are discussed and refined, the products and indicators might look something like this:

Product 1. Quality of repair.

Indicator A:	Return rate: percentage of items repaired that were returned for reaccomplishment of repair.
Indicator B:	Percentage of Quality Control inspections passed.

Product 2. Ability to meet demand for repairing items.

Indicator:	Number of units repaired divided by total number of units brought in for repair.

Product 3. Ability to meet training needs.

Indicator:	Number of people qualified to work on each type of item to be repaired, divided by number of people needed to be qualified.

As mentioned above, this would not be a complete list of products and indicators for such an organization, but it does serve to explain the concept of the productivity measurement system.

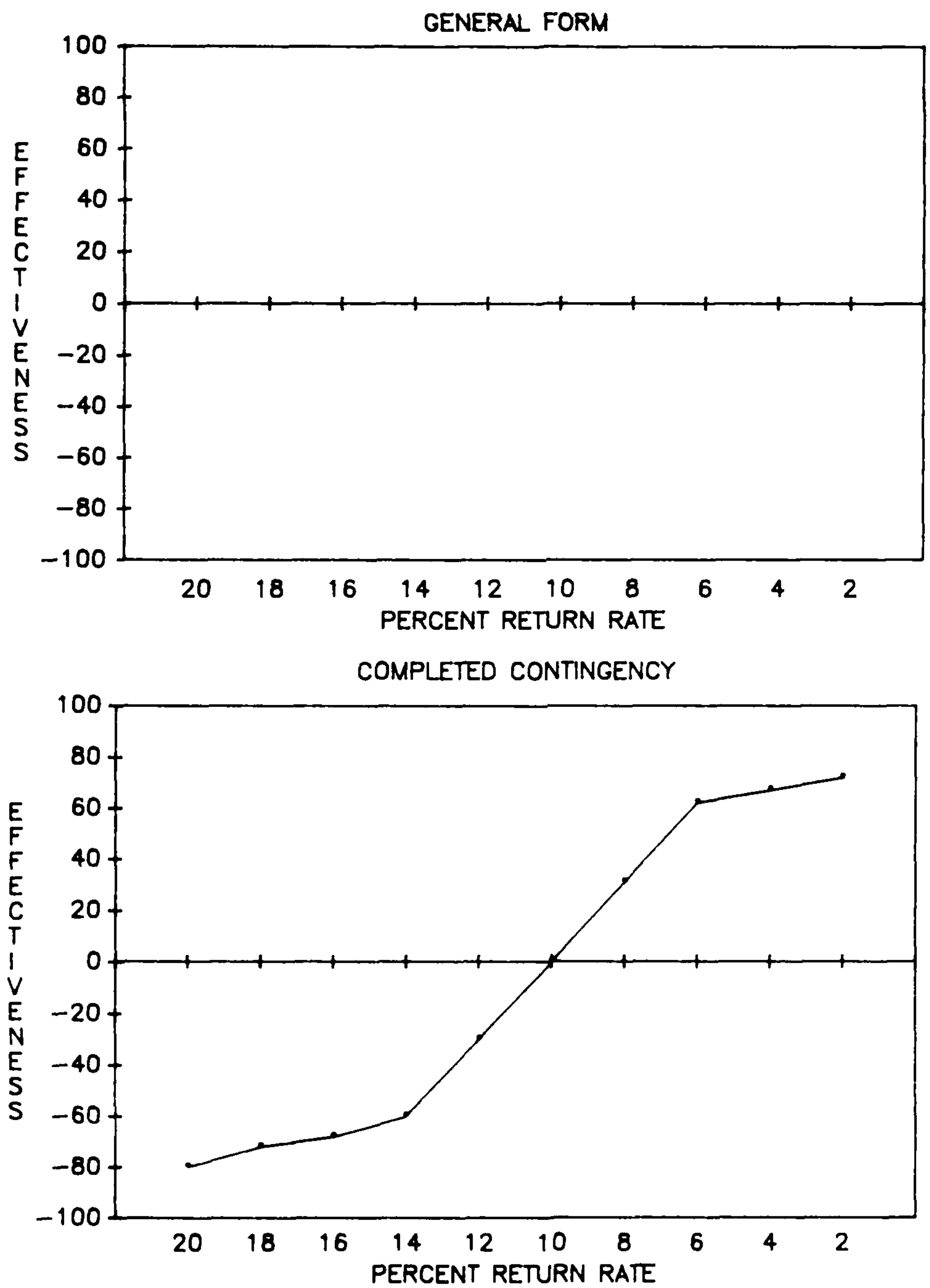
Step 3. Establish Contingencies.

Once the products and indicators have been identified and agreed upon, the next step is to establish the contingencies. A *contingency* is the relationship between the amount of the indicator and the effectiveness of that amount of the indicator. Figure 1 presents an example of a contingency. The top half of the figure shows the general form of a contingency. The horizontal axis represents the amount of the indicator, ranging from the worst possible level to the best possible level. For this example, we have chosen the first indicator: the percentage of items returned for reaccomplishment of repair. Assume that the people in the organization have said that the best possible return rate is 2% because about 2% of the electronic components they use for repairs can work properly when installed and checked, but fail almost immediately when put into use. Let us also assume they said that the worst possible return rate would be 20%. Based on this information, values on the horizontal axis would fall between 2% and 20%. The vertical axis of the figure shows the effectiveness values of the various levels of the indicator. It ranges from +100 which is maximum effectiveness, to -100, minimum effectiveness. It also has a zero point which is defined as the expected or neutral level of effectiveness. That is, the zero point is neither positive nor negative.

Once the best and worst possible levels of productivity have been established by the organizational personnel, the next task is to determine the zero point; that is, the indicator's expected level, the level that is neither especially good nor especially bad in terms of productivity. Once this is established, a point would be placed on the figure at the intersection of the zero point of the vertical axis and the level of neutral point on the horizontal axis. For example, if the neutral point were identified as a return rate of 10%, it would be indicated as shown in the bottom half of Figure 1.

Next, the maximum and minimum effectiveness levels would be established. The first step is to list the maximums for each of the indicators. Assume the maximum indicator levels for the four indicators in our example were as follows.

FIGURE 1. EXAMPLE CONTINGENCY



Indicator	Maximum Possible Value
1. Percent return rate	2%
2. Percent quality inspections passed	100%
3. Percent repair demand met	100%
4. Percent qualified/needed	130%

The group of incumbents and supervisors is then asked to rank order these maximums in terms of the contribution of each to the overall effectiveness of the unit. The one that the unit personnel believe to be the most important thing that the unit can do is given a rank of 1, the second most important thing is given a rank of 2, etc. The group discusses this and consensus is reached. The maximum with the highest importance rank is then given an effectiveness value of +100, and the group is asked to rate the other maximums relative to this. They are told to rate the other maximums as percentages of the +100 maximum. For example, if the maximum of a given indicator was only half as important to the effectiveness of the unit as the most important maximum, they would give it a value of +50. An analogous process is then done for the minimum values of each indicator, except the most important (worst) minimum is not constrained to be a value of -100, it is given the value that the group feels is appropriate.

Once the zero points are identified and the effectiveness values of the maximums and minimums established, the remainder of the points in the line are developed by the group. Group discussion is continued until consensus is reached.

Assume that the personnel in the organization said that return rate was an important aspect of their work, and that to be at the minimum (20% return rate) would correspond to an effectiveness of -80, and to be at the maximum (only 2% return rate) would be a +70. After the other points on the line were identified, this might result in a contingency similar to that shown in the bottom of the figure, which indicates that going above the neutral point results in increasing positive values, but such are not linear. In the example, once a return rate of 6% is reached, lower return rates do not represent as great an increase in effectiveness. Likewise, at the low end, once the return rate reaches 14%, they are doing very badly, and any rate below that is proportionally not as bad.

After this process has been completed for each of the indicators, and once all have been scaled and reviewed for accuracy, the contingency set would be complete.

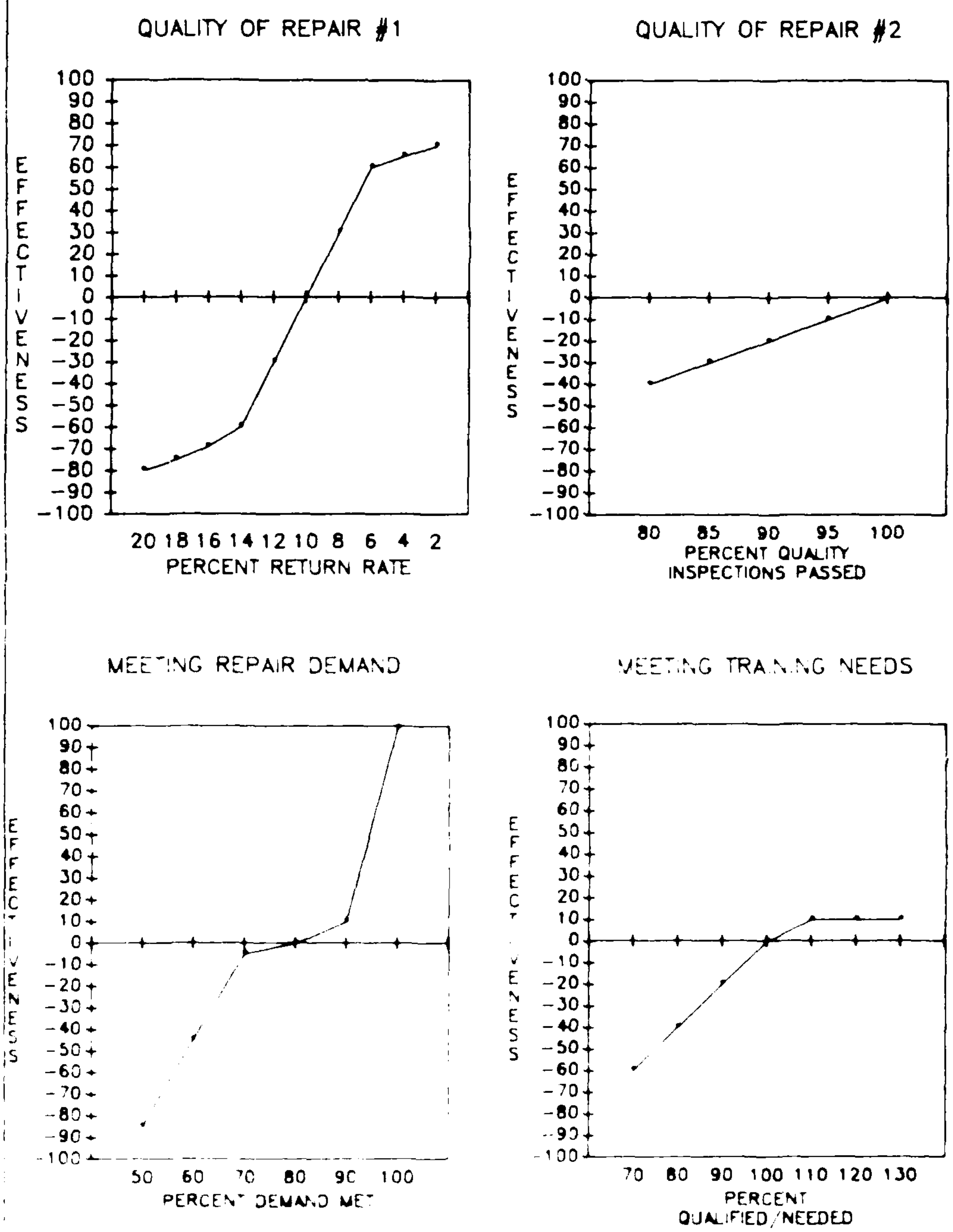
A sample contingency set is presented in Figure 2. For each indicator there is a contingency with its maximum, minimum, and expected or zero point, and a function relating it to effectiveness. The first contingency, that for return rate, is the same as that shown in the lower part of Figure 1. The second contingency is for the percent of Quality Control inspections passed. Note that for this contingency the expected level is that 100% of these inspections be passed. Recall that these inspections are not inspections of the final work but rather, inspections of the process the technician goes through in doing the repair. It is an index of how well the person is following the manual in doing the repairs. It is expected that all repairs will be done in accordance with the manual. Thus, this contingency shows that the expected level is doing all repairs (100%) in accordance with the manual. Anything less than this is below expectations, and results in negative effectiveness. Note that in this particular case (since it is not possible to pass more than 100% of inspections), there are no positive values. Together, these two contingencies cover Product 1, Quality of Repair.

Product 2, Ability to Meet Demand for Repairing Items, has only one indicator; and hence, only one contingency. The indicator is the number of units repaired divided by the total number brought in, expressed as a percentage. This contingency is steep at the low and high end, and fairly flat in the middle section.

Product 3, Ability to Meet Training Needs, also has only one indicator: number of people qualified (through training) to repair equipment, divided by the number needed, expressed as a percentage. For this indicator, it is possible to go above 100% qualified since, although the organization needs only 15 people to be qualified to repair a given piece of equipment, it could actually have more than 15. However, the contingency becomes flat after 110%, indicating that having more than 110% is no more effective than having 110%. The idea is that once there is a small excess over the maximum number needed, having additional trained personnel is not important.

Two things are particularly noteworthy about these contingencies. The first is that the overall slope of the function expresses the relative importance of the indicator. For example, the overall slope for the first indicator (return rate) is steeper than that for the second indicator (percent of inspections passed). This reflects the fact that although it is important to pass inspections, which show that the process of doing the repair was accurate, actually doing the repair so that the item functions properly is more important. Second, the contingencies can be non-linear. As shall be discussed below, this is necessary to accurately reflect the realities of an organization's functioning. In many cases, the relationships that actually exist are simply not linear.

FIGURE 2. EXAMPLE CONTINGENCY SET



It is important to recognize what these properties of the contingencies do. Many productivity measurement systems, even if they attempt to measure all the important aspects of the organization and combine them into a single index, do so by some form of summing of the measures. This amounts to saying that all the functions of the organization are equally important. Clearly, this does not reflect organizational reality. Different things the organization does are not equally important. Our system deals with this differential importance issue by the nature of the contingencies. Aspects of the work that are very important get steeper contingencies than aspects that are less important. For example, in Figure 2, the first indicator for Quality of Repair has a range in effectiveness from -80 to +70, whereas Ability to Meet Training Needs ranges from -60 to +10. This indicates that Quality of Repair is more important than Ability to Meet Training Needs, since variations in Quality of Repair have a greater impact on the effectiveness of the organization. Thus, the *relative* importance of each different aspect of the work is incorporated into the contingencies.

Another approach that could be used to incorporate differential importance into the measurement system is to measure each aspect of the work and then weight each measure according to its importance. (Presumably each would be first divided by its standard deviation to equalize the relative contribution before weighting for importance.) Thus, for example, Quality of Repair might be considered as being twice as important as Ability to Meet Training Needs. To get an overall productivity index, the Quality of Repair measure would be multiplied by 2 and added to the training measure. We feel our approach is superior to this technique. The problem is that the simple weighting method assumes that there is a linear relationship between amount of the measure and productivity; that is, to improve a given amount at the low end of the measure is as good as improving that same amount at the high end. However, in the real world, it is very common for values in the middle range of an indicator to represent large improvements in productivity, and values at the high end to represent a point of diminishing returns. That is, once an organization gets to a fairly high level of productivity on one aspect of the work, it is frequently better to try to improve something else that they are not doing as well, rather than continue to improve something that is already at a high level.

For example, if the repair shop were operating with a very low return rate, it might be better to try to improve meeting its training needs rather than attempting to further improve its return rate. Thus, even though return rate overall is more important than training, if return rate is good, improving a low degree of training readiness can become more important to the overall effectiveness of the organization. Another example of this non-linearity would be a situation such as that depicted for training needs. Once the organization reaches a certain point, further increases are not more effective since all the people that are necessary are already trained.

The simple weighting method ignores this non-linearity because no matter where the organization is on the measure, the value is always weighted by a constant amount. The contingencies in our system capture this non-linearity and thus provide a more accurate picture of the organization's functioning.

Step 4. Put the System Together.

Once the contingencies are completed and approved by management, the last step is to put the system together. This would be accomplished by first collecting the indicator data for a given period of time. If the time period selected were a month-long period, the data for the four indicators would be collected at the end of the month. Then, based on the contingencies, effectiveness scores would be determined for each indicator by calculating the effectiveness for that level of the indicator. This is illustrated in Figure 3. For example, if the maintenance unit had a return rate of 6% in the month of March, examining the contingency indicates that such a return rate is associated with an effectiveness score of +60 (i.e., a value of 6% return rate on the horizontal axis is associated with an effectiveness value of +60 on the vertical axis). Continuing this process would give an effectiveness value for each indicator, as exemplified in Figure 3.

Once the effectiveness values are determined, they can be summed to derive the overall effectiveness score for products with more than one indicator, as seen for the first product. The total effectiveness of the product Quality of Repair would be the sum of the two indicators comprising that product: +60 for Return Rate and -10 for Percent Quality Control Inspections Passed, for a total of +50. Next, overall productivity can be calculated by summing all of the effectiveness scores. In the example, this Overall Effectiveness score is +20.

These effectiveness scores have a distinct meaning, in that a score of zero means that the organization is meeting expectations; that is, their productivity is neither particularly good nor bad. As the score becomes positive, they are exceeding expectations. The more positive the score, the more they are exceeding expectations. As the score becomes negative, they are below expectations. The closer they are to the maximum possible overall effectiveness score, the closer they are to their best possible productivity.

This ability to simply sum effectiveness scores is one of the major advantages of the system. Because the contingencies reflect the relative importance

Figure 3. Completed System.

PRODUCTIVITY: MAINTENANCE UNIT

DATE: March, 1987

	INDICATOR DATA: MARCH	EFFECTIVENESS SCORE
I. Quality of Repair		
A. Return Rate	6%	+60
B. Percent Quality Control Inspections Passed	95%	-10

Total Effectiveness: Quality of Repair = +50

II. Meeting Repair Demand		
A. Percent Demand Met	90%	+10
III. Meeting Training Needs		
A. Percent Qualified/Needed	80%	-40

OVERALL EFFECTIVENESS = +20

and the non-linearity of the indicators, these factors are already incorporated in the system; thus, a simple summing does indeed reflect the overall effectiveness of the unit. As will be discussed later, this property also makes it possible for the system to be used to aggregate across individual units to determine the productivity of larger and larger units of the organization.

Accuracy Of The System

In order for this approach to be a good measure of organizational productivity, it is obvious that it needs to be accurate. This means several things. It means that the listing of products and indicators must be complete. If there are important functions of the unit that are not included among the products, or if important indicators are omitted, the system can easily produce a situation where those things that are measured are attended to, and those that are not measured are somewhat ignored. This uneven attention to important functions can have very dysfunctional consequences for the organization. A second aspect of accuracy deals with the degree to which the system must accurately reflect what the unit *should* be doing. This means that the products, indicators, and contingencies must also be correct.

Both completeness and accuracy are dealt with in the development of the system by having a clear process of approval of the system at higher levels of the organization. This approval process is made clear from the start. That is, at the beginning of the development of the system, all participants are told that incumbents and supervisors will develop the products and indicators, which will then be presented to higher management for approval. Once higher management has approved the products and indicators, the supervisory groups develop the contingencies, which must also be formally approved by higher management. While this approval mechanism, and the multiple inputs that it provides, does not guarantee completeness and accuracy of the system, it provides a system of checks and balances so that the system will be as complete and accurate as possible.

A final point about the quality of the system is that the development of the system necessarily introduces subjectivity into the system. Subjectivity is present in the listing of the products and indicators, and especially in the ratings that are used in the contingencies. Subjectivity is present, but this is not necessarily a problem. The elements of the system--products, indicators, and contingencies--are in actuality statements of policy. As a whole, they say (a) what is important to the functioning of the unit, (b) the level of output that is expected (the zero point), (c) how good other levels of output are, and (d) the relative importance of different types of functions for the unit. These determinations represent policy, and policy is a subjective thing. A manager's primary responsibility is to set policy, in the sense that he/she must determine the priorities for resource allocation within his/her unit. What our system does is to reduce ambiguity in policy and priorities by formally discussing them, quantifying them, and subjecting them to formal review and approval by the management of the organization.

Priorities

There are two other unique features of the productivity measurement system that should be described; the first is the system's capability to generate unit priorities.

The system offers a way to develop a clear set of priorities for improving productivity. Recall from Figure 3 that for a given time period (e.g., a month), the system presents the actual amount of each indicator achieved for that period, and the effectiveness levels of those amounts of the indicators. It would be a simple matter to look at the contingency for each indicator and calculate the effectiveness gain that would occur if the unit went up one increment on each of the indicators during the next period. For example, if the unit had a Return Rate of 6% in March, as is indicated in Figure 3, for them to go to the next level up (a 4% Return Rate) in April would mean an increase in effectiveness from +60 to +65, for a gain in effectiveness of +5 units. This could be calculated for each indicator. Once it was calculated, one could rank order the changes from highest to lowest. Such a listing for our example would look like that in Figure 4.

This information communicates exactly what should be changed to maximize productivity. In the example it says that the best thing the unit can do is to increase their meeting of repair demand. That is where they should focus their efforts if they want to best increase their productivity. Once this is done, or if increasing on this factor is not possible, the next best thing they could do is to improve training so that more people are qualified. Improving return rate and improving quality control inspections are the least important in increasing productivity, with improving on quality control inspections being slightly more important than improving return rate.

Thus, the system can generate a set of priorities that unit personnel can use to guide efforts to increase productivity. This would aid in decisions about resource allocation, and where to focus to identify barriers to productivity.

Figure 4. Priorities for Increasing Productivity.

PRIORITIES FOR: APRIL, 1987

CHANGE	GAIN IN EFFECTIVENESS
Percent Demand Met from 90% to 100%	+90
Percent Qualified (Training) from 80% to 90%	+20
Percent Quality Control Inspections Passed from 95% to 100%	+10
Return Rate from 6% to 4%	+5

Aggregation Across Units

A second unique feature of the productivity measurement system is the ability to aggregate across organizational units. It is quite valuable to have a productivity measurement system for a given unit, or several units. It would be even more valuable if one could aggregate the measurement system from the several different units into one measure that indicates the total productivity across all the units. For example, if a branch were composed of several separate sections, it would be valuable to have a measure for each section, and be able to combine those section measures into a single measure for the entire branch. In most productivity measurement systems this is not possible, since the measurements vary from unit to unit. An advantage of our approach is that it is possible to do such across-unit aggregation. Each unit is measured on a common metric: overall effectiveness. Since each of the sections is measured on this common metric, it becomes possible to simply add the overall effectiveness of each of the sections to get a measure of the overall effectiveness of the branch, as long as the scaling of the contingencies is done with this aggregation in mind.

If one were to simply add the overall effectiveness scores of the different sections to determine the productivity of the branch, one would be essentially assuming that each section contributes equally to the effectiveness of the branch. Although this may indeed be the case, it is not safe to assume it. It could easily be that the work of one unit is more critical than that of the others, and thus, this unit's effectiveness would contribute more to overall effectiveness than that of the others. Another likely possibility is that a section with 40 people is going to make more of a contribution to the organization than a section with 5 people.

Dealing with this problem is actually a fairly straightforward matter. When the system is developed for a single section and all levels of supervision and management have agreed on the values, we assume that the contingencies are accurate for that unit. That is, because all personnel from incumbents to senior managers have agreed on the contingencies, they are accurate reflections of policy. Specifically, we assume that the values for each contingency are accurate relative to the other values for that contingency, and that all the contingencies for the given section are accurate relative to the other contingencies for that section. After the contingency set for each unit has been developed, all that remains to be done is to rescale the contingency sets for accuracy across sections.

To explain this rescaling across sections, let us use an example of an Avionics Branch in a Component Maintenance Division. Assume that this branch has three sections. One section is the electronic communication maintenance section we have been using as the example; we will call this unit

Communications. The branch also has two other sections, one doing maintenance on Inertial Navigation, and one maintaining Automatic Flight Control Instruments.

To do the rescaling, we start with the fact the most important indicator for each section will have an effectiveness value of +100. This is true simply because, in the development of the contingencies, the most important indicator is always defined to have an effectiveness value of +100. Thus, by definition, each of the sections in the same branch will have at least one indicator value of +100. In doing the actual rescaling, we take the indicator with the +100 effectiveness score for each section. This can be thought of as the most important indicator for each section. With the three sections in this branch, there would be three such indicators (i.e., one for each section).

The top indicators from the three sections would then be shown to branch management, as well as to managers from levels of supervision above the branch. The managers would then be asked as a group to rescale the three levels. To do this, they first rank the three levels in terms of overall contribution to the branch. That is, they are asked which of the three outcomes they would most value for the overall effectiveness of the branch. In the example, they are given the three levels: 100% repair demand met in Communications, 100% in Inertial Navigation, and 100% in Automatic Flight Control Instruments. They are then asked to indicate which of the three outcomes would make the greatest contribution to the effectiveness of the Avionics Branch. They discuss this and come to a consensus. It could be that they believe that since all three types of components are crucial for an aircraft to be operational, they are all equally important. In contrast, they may feel that the three are not equally important. This could come about for a variety of reasons. For example, one section might repair more items and have a correspondingly larger number of personnel. Thus, that section makes a larger contribution to the overall functioning of the branch than the other sections. Differential importance could also occur because there is a sufficient backlog of repaired components in supply for two of the sections, but not for the third. Thus, repair of the non-backlogged components is more crucial for meeting mission requirements than repair of components for which there is already a sufficient backlog.

Assume that in the discussions, unit supervisors and managers decide that meeting 100% repair demand in Communications is the most important thing for the branch, meeting 100% repair demand in Automatic Flight Control Instruments is next, and meeting 100% repair demand in Inertial Navigation is next.

Once this ranking is completed, the indicator ranked highest in effectiveness is given a value of +100, and managers are then are asked to rate the remaining indicators relative to this one. In doing this, they are told to think in terms of percentages; that is, they should ask themselves whether the second most important thing is 95% as important as the most important, 90%, etc. Assume that the second most important maximum is given a value of 90, and the third a value of 75. This means that the group is saying that meeting 100% repair demand in Automatic Flight Control Instruments is almost as important as meeting 100% repair demand in Communications, but not quite as important. In fact, they are saying that it is 10% less important. Meeting 100% repair demand in Inertial Navigation is somewhat less important, and in fact is only 3/4 as important as meeting demand in Communications.

Once the relative ratings of the top indicators for the sections are agreed upon, the next step is to rescale the individual contingencies for each of the sections. This is done by reducing the effectiveness score of each level of each indicator in a given section by the percentage its own maximum indicator was reduced in the rescaling. For example, in Automatic Flight Control Instruments the original +100 maximum was reduced to +90 in the rescaling process. This represents a decrease of 10%. In essence, it is saying that to be comparable with the other sections in the branch, the effectiveness of this level of the indicator must be reduced by 10%, since it is not quite as important to the branch as the maximum of the Communications section. Since the effectiveness value of the maximum was reduced by 10%, in order to retain accuracy it is necessary that the effectiveness levels of <u>each</u> of the indicators for that section also be reduced by 10%. This means that if the original positive values of one of the indicators for that section were +10, +20, +40, and +75, the values after the rescaling process would be +9, +18, +36, and +67.5. This process of reduction by 10% is continued for each contingency in that section.

A similar process is then done for the Inertial Navigation section. Here the maximum was reduced by 25%, from +100 to +75. Thus, each positive effectiveness score for each contingency in that section must be reduced by 25%.

Finally, the effectiveness values for the Communications section do not change. Since the original maximum of +100 was unchanged in the process of rescaling across sections, the effectiveness values for the contingencies in Communications are not recalculated.

An analogous rescaling process is then done for the negative effectiveness values of the indicators. The most negative level of the indicators is listed for each of the branch's three sections. These three levels are then ranked as to which constitutes the poorest level of effectiveness. Finally, just as with the positive values, the negative values of each level of the indicators are each adjusted by the percentage that the original minimum indicator level was reduced.

This rescaling process has the effect of adjusting the effectiveness scores of the different sections in the branch for any differences in importance of the different sections. Once it is finished, the overall effectiveness values from the different sections are calculated just as before. The only difference is that the effectiveness scores for the different levels of the indicators have been rescaled based on the aggregation process. The overall effectiveness score from each of the sections can now simply be summed to produce overall effectiveness of the entire branch. For example, if the monthly overall effectiveness for Communications was +250, for Internal Navigation was +150, and for Automatic Flight Control Instruments was +200, the total branch effectiveness would be +600. This value can be interpreted just like overall effectiveness for a single section. If it is 0, the branch overall is meeting expectations. If it is above 0, the branch is exceeding expectations and the higher it is above 0, the greater they are exceeding expectations.

This approach to aggregation can be extended to larger and larger units of the organization, so that, if desired, a single index of the productivity of the entire organization can be developed. For example, one could aggregate branches into a division index. Assume that in the Component Maintenance Division that we have been discussing, there is not only an Avionics branch like that described above, but there is also a second branch called Propulsion. This branch has two sections, Jet Engines and Test Unit. Jet Engines repairs and rebuilds engines. Test Unit runs the repaired/rebuilt engines in a test facility to evaluate the functioning of the engine.

To aggregate this second branch up to the level of the division, the first steps are to develop the products, indicators, and contingencies for the two sections in the Propulsion Branch. Once this is done and the system for each section is approved up the chain of command, the next step is to aggregate the two sections into the measure for the Propulsion Branch, and to aggregate the two branches into a measure for the Component Maintenance Division.

To achieve these two levels of aggregation (section to branch and branch to division) is fairly straightforward. In essence, we do the same thing that was done for the Avionics Branch. Instead of rescaling the three maximums from the three sections in Avionics, we rescale the *five* maximums from the *five* sections in the Component Maintenance Division at the same time. Put another way, if one wants to aggregate up to the level of the Division, the aggregation to the branch and to the division is done at the same time, and by the same process.

Specifically, we would take the maximum indicator value from the five sections and list them, just as we did when aggregating the three sections of Avionics to the branch index. This would result in a list of the five levels of the indicators that received the +100 effectiveness value. For the three sections in Avionics, this would be meeting 100% of repair demand in the three sections of Communications, Inertial Navigation, and Automatic Flight Control Instruments. Added to these three would be the indicator levels in the two sections in the Propulsion branch that had received the +100 effectiveness scores. Assume that for the Jet Engines section the indicator level with the maximum effectiveness value (+100) was having 6 or more jet engines repaired, inspected, and ready for installation. Assume that for the Test Unit section, the indicator level with the +100 effectiveness score was having 0% engines that had been passed by the section returned as malfunctioning.

These five maximum indicator values would be ranked and rated just as in the example of using only the three maximums from the Avionics branch. Assume that the ratings came out as follows.

Section	Maximum	Rating
Communications	100% Repair demand met	100
Jet Engines	6 or more engines ready for installation	98
Test Unit	0% engines returned as malfunctioning	95
Auto Flight Control	100% Repair demand met	90
Inertial Navigation	100% Repair demand met	75

In other words, the indicators from the Propulsion Branch were seen by the group as slightly less important to the functioning of the division than the most important maximum from the Avionics Branch, but more important than the other maximums from the Avionics Branch. Once these values have been determined, the final step is to recalculate the effectiveness values of the indicators, as was done in the previous example. Effectiveness scores for the Communications section would remain unchanged since the original maximum with its value of +100 is still +100 after rescaling. The effectiveness values for the indicators in the other four sections change. Each positive value in Jet Engines is reduced 2%, 5% in Test Unit, 10% in Auto Flight Control, and 25% in Inertial Navigation.

As before, an analogous process is done with the indicator values with negative effectiveness scores. The five negatives are ranked and rated by the group and then the negative effectiveness scores are rescaled by the percentage that the maximum for that section was reduced.

Once the rescaling of the indicator values is completed, overall effectiveness for each section can be calculated as usual, using the rescaled effectiveness scores. Branch overall effectiveness is simply the sum of the section overall effectiveness scores, and division overall effectiveness is simply the sum of the two branch overall effectiveness scores. By going through this process of rescaling, the sections in each branch and the branches in the division are made comparable to each other. This simple summing of overall effectiveness scores preserves the relative importance of the different sections and branches.

This same logic of rescaling to make the different units comparable with each other can be continued to larger and larger units. In theory, it is possible to use this approach to develop a single index of productivity for the entire Air Force.

One potential problem that could come up in this rescaling process is that the shape of the original contingency could change. Recall that rescaling is done on the positive effectiveness values, then repeated for the negative effectiveness values. In other words, the rescaled effectiveness values for the best possible levels of the indicators are determined, then the process is repeated for rescaling the effectiveness values for the worst possible levels of the indicators.

This two step process could have the effect of changing the shape of the contingency that was originally developed by the unit and its supervision. For example, suppose a given section developed a contingency that was linear. That is, the contingency was a straight line from the worst level of the indicator (with e.g. a -75 effectiveness value) to the best level of the indicator (with a +75 effectiveness value). After rescaling, the maximum could stay at +75, while the minimum was rescaled to -50. This would mean that the rescaled contingency was no longer linear in shape. It would be steeper from the zero point to +75 than it would be from the zero point to -50.

This in and of itself is not a problem. The new non-linear contingency would be the most accurate reflection of the contribution of amounts of that indicator to overall organizational effectiveness, and the new contingency would be used in calculating effectiveness values for unit feedback reports. The potential problem is that the unit personnel developed a contingency that has now been changed by the aggregation process. This could lead to a lack of acceptance of the new contingency by unit personnel.[1] Thus, it is important to explain to unit personnel from the start what the aggregation process will do and how it could change the contingencies. In addition, if contingencies are changed, the reasons for these changes should be explained to unit personnel by the unit supervisors who were present at the meetings that did the rescaling.

One final point should be made about this rescaling process to enable aggregation to larger units to be done. The description of the rescaling process sounds rather complicated, but its implementation is really quite simple. It only requires that the appropriate supervisors and managers be brought together for one or two meetings to rank and rate the maximums for the sections to be aggregated. While this can be a difficult set of judgments for them to make, it takes only a short amount of time. This is especially true since at this point in the process, these personnel have been involved in the development of the system for some time and should be quite familiar with the issues. The longest we would expect such meetings to take would be two hours. Once the judgments are made, it is a simple matter to recalculate the effectiveness scores. Once this is done, the rescaling is finished.

IV. APPLICATION OF THE SYSTEM

Although we believed this approach to measuring productivity to be conceptually sound, it remained to be seen if it would actually be successful in an Air Force environment. To explore this, a field test of the system was undertaken as part of an AFHRL-funded effort conducted by the authors at an operational Air Force base in the Southwestern United States. In the overall productivity project we focused on the interventions of feedback, goal setting, and incentives as techniques for enhancing organizational productivity. The productivity measurement system described here served as the basis for the feedback, goal setting, and incentives, and was the criterion by which the three interventions were evaluated. The interventions and their results are presented in another AFHRL Technical Paper (Pritchard, Jones, Roth, Stuebing, & Ekeberg, 1987). Here we are focusing on the productivity measurement system itself.

The organizational units involved in the field test consisted of one maintenance section and four sections in resource management. The maintenance unit was the Communications and Navigation section (Comm/Nav) in the Component Repair Squadron. The four sections in Resource Management, which together comprised the Materiel Storage and Distribution Branch (MS&D) of the Supply Squadron, were Receiving, Storage and Issue, Pickup and Delivery, and Inspection.

[1]Aggregation techniques which preserve the functional form of the original contingencies are theoretically possible but have not been fully explored as part of this effort. Additional work on this topic is planned at AFHRL.

The Comm/Nav section was similar to the example of the electronics maintenance unit used above in describing the development of the system. Its function was to repair a variety of electronic equipment used for aircraft communication and navigation. The MS&D branch was essentially the base warehouse. Property was delivered to the warehouse and checked in by the Receiving section; Storage and Issue shelved the property and retrieved it as it was ordered by units on the base; the Delivery section had the responsibility for delivering the property to units on base that had ordered it; and Inspection's responsibility was to make sure the property was in good condition, and ensure that regulations were being followed concerning packaging, storage, and identification of property.

System Development

Once the five units for study were selected, the first step in the development of the system was to meet with supervision and management of the units and explain the purpose of the project and what the research team would be doing. The development of the productivity measurement system was also explained, and questions were answered. After this, meetings were held jointly with incumbents and supervisors to actually develop the system. Most of the system development was done through group meetings. These meetings were held on visits to the base made approximately every two weeks over a period of six to nine months.

Once the missions of the units were examined, the first step was to identify the products of each of the units. This task proved to be a fairly time-consuming one. There was considerable debate on exactly what the important products of the units were. We continued these meetings on products until consensus was reached.

The same process was used in the identification of the indicators. The meetings were held over a considerable period of time since there was a great deal of discussion and initial disagreement as to what indicators would be appropriate. During this process, it became clear that examining the actual data from various proposed indicators would be of value. This process was very instructive to the unit personnel. They realized that some existing measures were in fact measuring quite different things than they had thought, and they found measures that they did not know were available. It also became clear that a number of new measures would have to be collected to get a set of indicators that would adequately represent all the important functions of the unit. As with the development of the products, the meetings were held approximately every two weeks over a period of three to six months.

One issue that surfaced with the MS&D branch involved an indicator that was a very important one but could not be directly attributed to a single section. The indicator involved the number of "delinquent documents," warehouse property documents that, due to some error, were lost in the system. When this occurs, these documents must be searched out and the problem rectified. The problem with an individual document could be caused by any of the sections in the warehouse; this complicated the determination of which section should be responsible. At the same time, delinquent documents were very important to the operation of the warehouse, and needed to be included in the system.

The resolution of the issue was to have an indicator for the number of delinquent documents, but make it an indicator that applied only to branch productivity. This meant that the effectiveness score for the branch would be influenced by how well or how poorly the four sections did on this indicator; however, the effectiveness of the individual sections would not be affected. This approach proved successful, in that the number of delinquent documents was reduced substantially, and no individual section felt that it was being held accountable for another section's errors.

Once the products and indicators had been finalized by group consensus, the next step was to get approval of the products and indicators from higher-level management. This approval was obtained using the entire chain of command from the units up to and including the Deputy Commander for Maintenance for Comm/Nav, and the Deputy Commander for Resource Management for MS&D. These personnel were again briefed on the project, and given written versions of the proposed products and indicators. They were given time to study these documents, after which a meeting was scheduled for formal review. In this meeting, the products and indicators were presented, discussed in detail, and, after some revisions, approved.

The next step was the development of the contingencies for each indicator. Meetings were held with incumbents and supervisors of each unit. First, the maximum and minimum indicator levels were established; then, the zero points were generated. These decisions also took considerable time, resulted in considerable initial disagreement, and were arrived at over several meetings. Once consensus was established, the effectiveness scaling of the indicator values within each unit was started.

This scaling was started by listing the maximum possible value for each of the indicators for their unit. The group was then asked to rank order these maximums in terms of contribution to the overall effectiveness of the unit. The group finally reached consensus. The maximum value assigned the highest ranking was then given an effectiveness value of +100, and the group was then asked to rate the effectiveness values of the other maximums relative to this. They were told to rate the other maximums as percentages of the +100 maximum. For example, if the maximum of a given indicator was only half as important to the effectiveness of the unit as the most important maximum, they were told to give it a value of +50. This process was also done as a group by unit supervisors and incumbents, and continued until consensus was reached. An analogous process was done for the minimum values of each indicator, except that the least effective minimum was not constrained to a value of -100, but was given the value that the group felt was appropriate.

Aggregation Across Sections

Once the contingencies were developed in each section of MS&D, the scaling required to aggregate across the four sections to the branch level needed to be done. As described above, this involved first identifying the indicator value scaled as +100 in effectiveness in each section; i.e., the maximum indicator value for the most important or top indicator for each section. These are listed below.

Receiving:	Get Priority 2 property (the highest priority property other than extremely rare emergency aircraft parts) to Pickup and Delivery in an average of 10 minutes.
Storage and Issue:	Get Priority 2 property to Pickup and Delivery in an average of 10 minutes.
Pickup and Delivery:	Get Priority 2 property to customers on base in an average of 15 minutes.
Inspection:	Have 100% of the aircraft parts inspected by 4:00 PM each day.

These four indicator values were next ranked by supervision and management as to their importance to the overall functioning of the branch. The first three were ranked as equally important, and the fourth was ranked below these. The three number one ranks were then given an effectiveness value of +100; that is, their effectiveness values remained unchanged. The supervisors were next asked to rate the last indicator value. They felt its importance in terms of effectiveness was very close to the first three, and gave it a value of +98.

The positive effectiveness scores for each contingency were then reduced by the amount the original maximum had been reduced in the rescaling. Since the maximum for Inspection had been reduced from +100 to +98, each positive value of each contingency for this section was reduced by 2% of its originally scaled value. For example, an original scaling of +50 would be rescaled to +49. Because the positive effectiveness values remained unchanged for the other sections, the contingencies for these sections could be used intact.

By this process, the four sections were scaled so that effectiveness values could be directly compared across sections. In addition, it was now possible to sum the overall effectiveness of the four sections to get an overall effectiveness for the branch that reflected relative importance of the four sections. The rescaling in this case led to a very minor adjustment in what would have been the branch sums had the rescaling not been done. However, this says only that in this particular branch, the section maximums were regarded as very similar in their contribution to branch productivity. This will not always be the case, and the rescaling could result in major changes in values for the other units.

Management Approval Of System

Once consensus was reached on the contingencies, approval was obtained from higher management, using the same process as described above for the products and indicators. Many questions were raised by higher management to unit supervision in the meetings on contingencies, as well as in the meetings on products and indicators. Higher management asked for clarification of many points and wanted to hear the units' defense of their system. While the majority of the products, indicators, and contingencies were left as originally developed by the units, there were some changes made as a result of the approval meetings. The process was a positive one. The resulting discussion clarified the position of higher management to the units' supervisors, and the positions of the supervisors were made clearer to higher management. In all cases, the resulting changes represented compromises that seemed to satisfy both groups.

An example of a completed system of products and indicators is presented in Figure 5. This system is the one that was developed for the Comm/Nav maintenance unit. It served as the basis for the example used earlier in this paper; however, as can be seen, the actual system is more complex than the simplified example. The final system had 3 products and a total of 13 indicators. The products were Equipment Repair, which includes both quality and quantity measures; Training; and Other Duties, which includes several miscellaneous activities of varying importance. Figures 6 and 7 present examples of actual contingencies that were developed for the Comm/Nav unit. Figure 6 shows the

Figure 5. Comm/Nav Products and Indicators.

Product 1. Equipment Repair

Bounces: Percentage of repaired equipment that did not function immediately after installation.

Percent QA (Quality Assurance) inspections passed.

AWM: Number of units awaiting maintenance.

AWP: Number of units awaiting parts.

Demand Met: Percentage of equipment brought in for repairs that was actually repaired.

Product 2. Training

STS Tasks Completed: Mean number of standard (more basic) training tasks completed for personnel in training.

Percent Qual Tasks Completed, Comm: Mean percent of advanced training tasks completed for personnel repairing communications equipment.

Percent Qual Tasks Completed, Nav: Mean percent of advanced training tasks completed for personnel repairing navigation equipment.

Scheduled Training Tasks Overdue: Total number non-technical (e.g. military) training requirements not met on time for all shop personnel.

Product 3. Other Duties

Mobility Equipment: Number of pieces of equipment used for mobility exercises that were not calibrated by the shop on schedule.

PMEL Overdue: Number of pieces of shop calibration and test equipment that were not calibrated by the shop on schedule.

Percent 349 Errors: Percent of errors on a major manpower documentation form.

Missed Appointments: Number of formal on-base appointments missed.

FIG. 6 BOUNCES

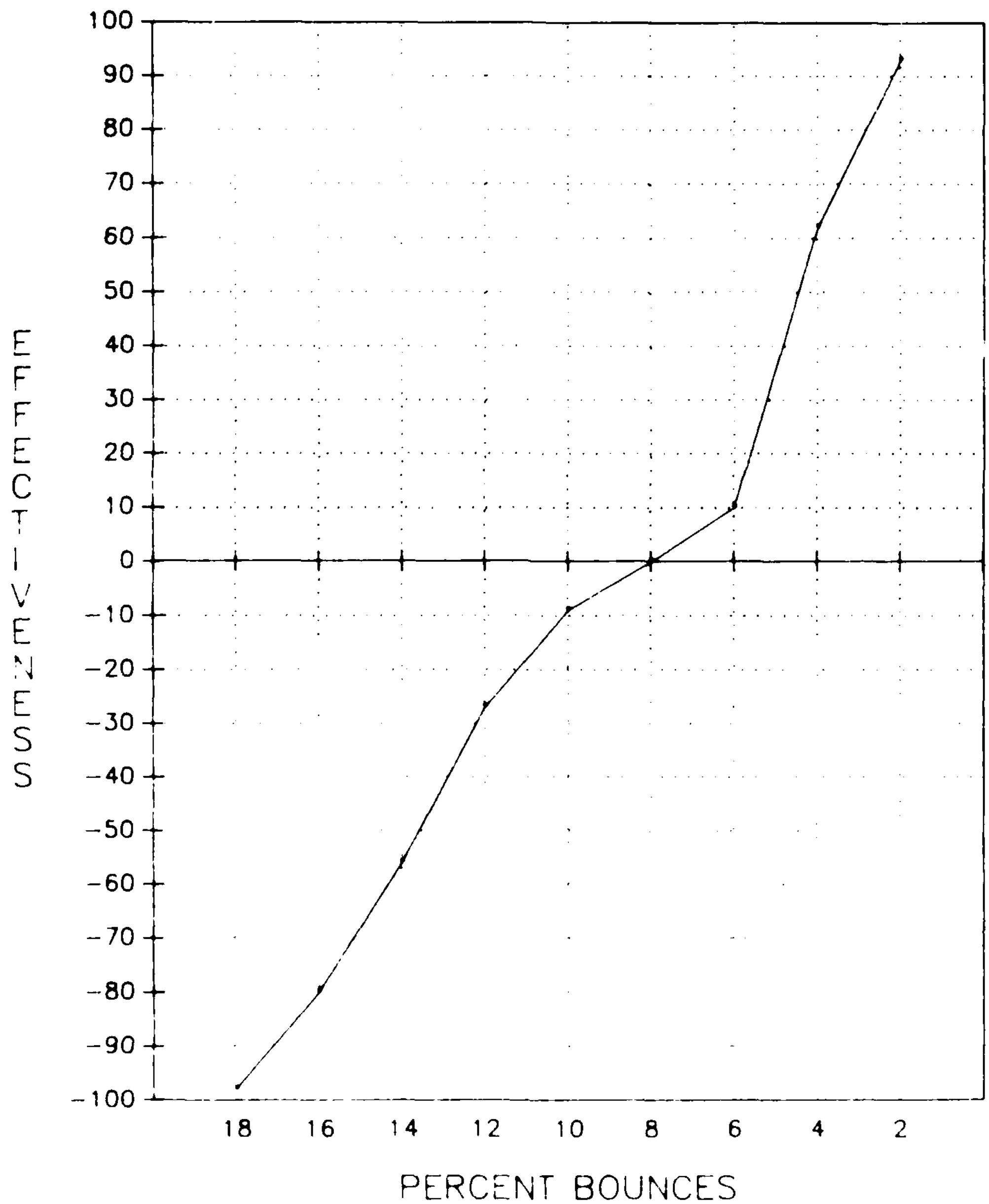
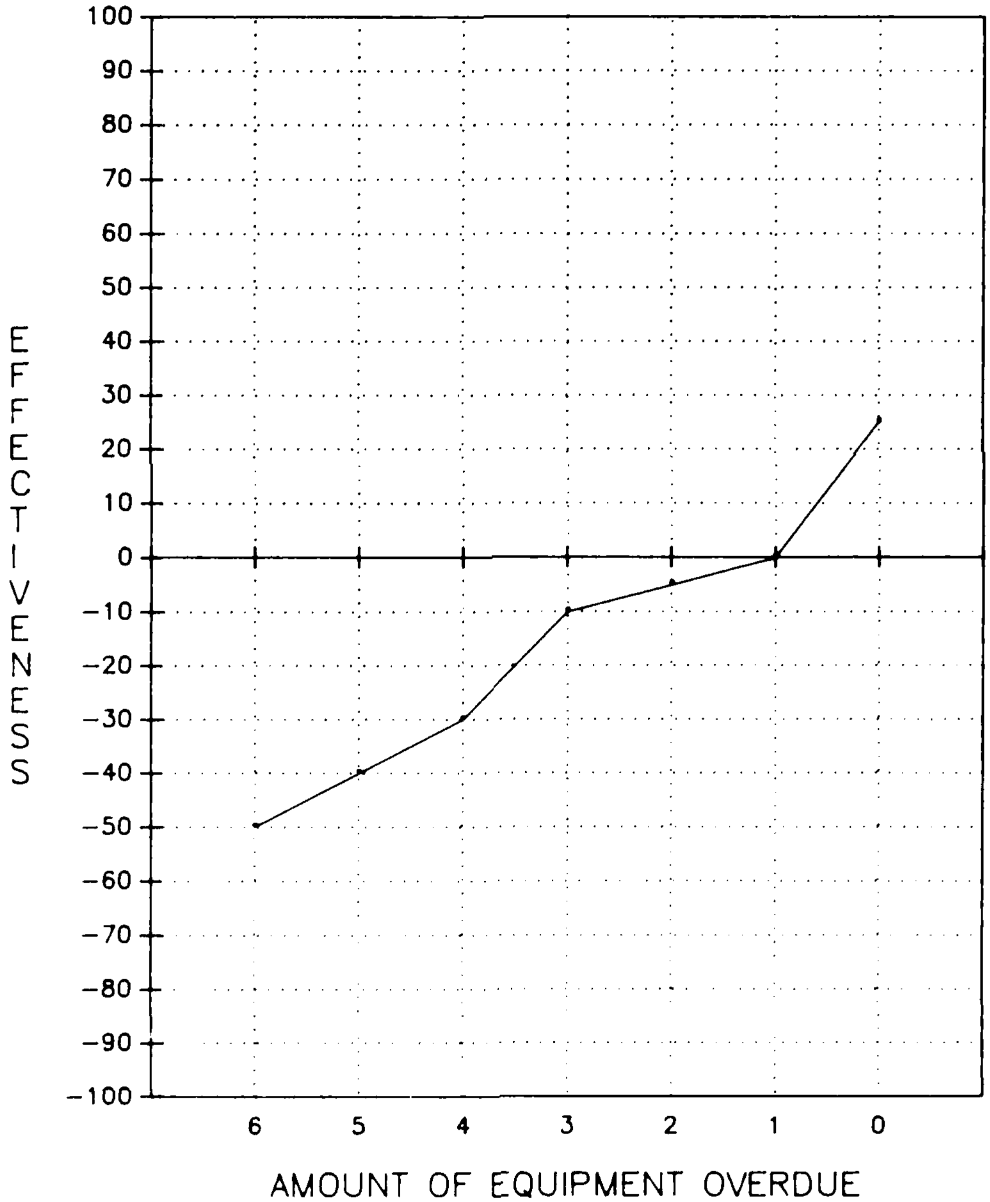


FIG.7 MOBILITY EQUIPMENT OVERDUE CALIBRATION



contingency for "Bounces" (i.e., the percentage of equipment where the repair needed to be reaccomplished). This is an indicator for Equipment Repair. It is a very important one, as shown by its range in effectiveness from +93 to -99. Figure 7 shows the contingency for Mobility Equipment Overdue Calibration. This activity is less important than Bounces, as shown by its more shallow slope, with its range from +25 to -50. The complete set of contingencies for all five of the units are presented in Appendix A.

Feedback

The final step was to put the system together to produce productivity feedback for each of the units. In doing this, we developed examples of what the basic system would look like and asked for management's thoughts on how best to present the material for maximum clarity. We also proposed some other information that they might find useful, and asked them for their suggestions on things to be included. This was discussed, and after several revisions, a final version of the productivity feedback report was developed. An example of the monthly report for the Comm/Nav shop is presented in Figure 8. Samples of the feedback reports for the four sections of MS&D are presented in Appendix B.

The first page of the report provides the basic productivity data. It shows the products and indicators, the indicator data for that month, and the effectiveness score associated with that level of each indicator. The lower portion of the page shows the total effectiveness for each of the products and finally the overall effectiveness for the shop. The second page of the report adds information to the basic data. The top half of the report shows the change in productivity from the previous month to the current month. The indicator data and effectiveness scores for both the previous month and the current month are shown, as are the changes in effectiveness from last month to the current month. This part of the report was requested by shop personnel to aid them in diagnosing areas where they were increasing or decreasing in productivity.

The bottom half of the page is the information on priorities for increasing productivity. For each indicator there is a column labeled FROM, TO, and GAIN. The FROM column is the amount of the indicator for the current month. The TO column is the amount of the indicator that represents an increase of one unit on the contingency and the GAIN column indicates the gain in effectiveness that would be achieved by such an increase. For example, for Demand Met, if the shop went from their March level of 91.7% to 95.2% in April, effectiveness would increase by 37 points. Examination of the GAIN column indicates that for the next month, the shop would increase their productivity most effectively by focusing on Quality Assurance Inspections, and the number of units Awaiting Parts

Figure 8. Sample Feedback Report.

PRODUCTIVITY REPORT
COMM/NAV SHOP
CRS MAINTENANCE

INDICATOR AND EFFECTIVENESS DATA FOR MARCH 1986

INDICATOR	INDICATOR DATA	EFFECTIVENESS SCORE
EQUIPMENT REPAIR		
BOUNCES	3.1	76
% QA INSPECTIONS PASSED	90.9	30
AWM	13.5	80
AWP	39.6	29
DEMAND MET	91.7	63
TRAINING		
STS TASKS COMPLETED	8	35
% QUAL TASKS COMPLETED: COM	69.5	72
% QUAL TASKS COMPLETED: NAV	56.8	68
SCHED TRAINING TASKS OVERDUE	0	10
OTHER DUTIES		
MOBILITY EQUIPMENT OVERDUE	0	25
PMEL OVERDUE	0	25
% 349 ERRORS	1	40
MISSED APPOINTMENTS	0	10

TOTALS	EFFECTIVENESS SCORE
EQUIPMENT REPAIR	278
TRAINING	185
OTHER DUTIES	100
OVERALL EFFECTIVENESS	563

Figure 8. (Concluded)

EFFECTIVENESS CHANGE FROM FEBRUARY TO MARCH

	INDICATOR DATA: FEBRUARY	EFF. SCORE	INDICATOR DATA: MARCH	EFF. SCORE	CHANGE IN EFF.
BOUNCES	2.8	81	3.1	76	-5
% QA INSPECTIONS	91.7	34	90.9	30	-4
AWM	15.58	72	13.5	80	8
AWP	40.6	27	39.6	29	2
DEMAND MET	91.5	59	91.7	63	4
STS TASKS COMPLETED	9	52	8	35	-17
%QUAL TASKS-COMM	68.6	72	69.5	72	0
%QUAL TASKS-NAV	59.5	71	56.8	68	-3
SCHEDULED TRAINING TASKS OVERDUE	0	10	0	10	0
MOBILITY EQUIPMENT OVERDUE	0	25	0	25	0
PMEL OVERDUE	0	25	0	25	0
% 349 ERRORS	2	27	1	40	13
MISSED APPOINTMENTS	0	10	0	10	0

CHANGE TOTALS

EQUIPMENT REPAIR	5
TRAINING	-20
OTHER DUTIES	13
OVERALL EFFECTIVENESS	-2

POTENTIAL EFFECTIVENESS GAINS FOR NEXT MONTH

	FROM	TO	GAIN
BOUNCES	3.1	0.4	17
QA INSPECTIONS	90.9	100	45
AWM	13.5	0	15
AWP	39.6	22.6	48
DEMAND MET	91.7	95.2	37
STS TASKS COMPLETED	8	9	17
%QUAL TASKS COMP: COMM	69.5	76	0
%QUAL TASKS COMP: NAV	56.8	62.8	4
SCHED TRNING TSKS OVERDUE	0	0	0
MOBILITY EQ OVERDUE	0	0	0
PMEL OVERDUE	0	0	0
349 ERRORS	1	0	15
MISSED APPOINTMENTS	0	0	0

(AWP). These show potential gains in effectiveness of 45 and 48 respectively. It would not be useful to devote attention to training in Comm Qualification Tasks, trying to further decrease Overdue Scheduled Training Tasks, or any of the other indicators that have a gain value of zero or near zero. This information can therefore serve as a basis for determining priorities for the next month. The unit should focus on those areas where the maximum gain in effectiveness could be produced.

The calculation of the GAIN amount is based on the amount of increase in effectiveness that would occur with an increase of "one unit" of the indicator. The size of a one unit increase was determined from the indicator values used in the contingencies. If the indicator values in a contingency were 2%, 4%, 6%, 8%, etc. the size of a one unit increase for that indicator was 2%. If the indicator values were 50, 60, 70, 80, etc., the size of a one unit increase was 10. The contingencies were originally developed so that the number of increments for the different contingencies was as equal as possible so that a "one unit" increment was roughly comparable across the different contingencies. Once the size of the "one unit" increase was determined for each contingency, the TO figure was calculated by adding the one unit increase to the actual value of the indicator for the preceding month. If the last month's indicator level was 83.6 and the size of one unit was 10, the FROM value would be 83.6, the TO value would be 93.6, and the GAIN value would be determined by what the contingency indicated as the gain in effectiveness if the unit went from 83.6 to 93.6 on that indicator.

There was one special circumstance that had to be dealt with using this approach. It was possible for the TO value to be higher than the maximum value of the indicator. This occurred when the unit was already high on that indicator and increasing "one unit" would put them over the maximum. It also occurred occasionally if the unit was already over the maximum on that indicator. This was dealt with by using the maximum possible effectiveness value for the indicator as the upper limit in effectiveness. In other words, if the effectiveness value for being at the maximum of the indicator was +75, this was the maximum effectiveness score that could be gained from that indicator. If the unit was near the maximum already with, for example, a past month's indicator level which yielded an effectiveness score of +73, the most they could improve would be to the value of the ceiling, +75, for a maximum gain of only +2.

The feedback report for MS&D was similar to this Comm/Nav report except that the report for each of the four sections included information on how the entire branch did for the month, including indicator and effectiveness data on the branch-level indicator discussed previously. In addition, the MS&D report added a feature that allowed for a direct comparison of the four units to each other.

One feature of the system is that it allows one to directly compare the productivity of very different units. This feature was very important to the supervisors and managers of the MS&D branch since it allowed them to compare the four sections of the branch. In order to make this comparison, we first determined the maximum possible overall effectiveness for each section. This was done by determining the effectiveness value for the maximum possible value of each indicator, and summing these effectiveness values. The resultant score represented the effectiveness value that would occur if the unit was doing as well as it was possible to do on every aspect of their work; in other words, their maximum possible effectiveness. Recall that these maximums were developed by consensus among the supervisors of the units, and discussed and approved by management. Thus, they should represent realistic maximums, and the effectiveness scores represent the value of the maximum contribution each of the units could make to the organization.

Once the maximum possible effectiveness was calculated, the actual monthly overall effectiveness score for each section was expressed as a percentage of maximum possible effectiveness. This percentage of maximum effectiveness was the measure by which each unit was compared to the other. These data were included in the monthly feedback report for each section of MS&D.

The feedback report was generated each month for each of the five units for the 15 months of formal intervention by the research team. It was presented within three workdays after the end of the month, and a copy was given to each individual in the chain of command, from the section supervisors to the Deputy Commander. A copy was also posted in the working area of each section so that incumbents could review it. In addition, graphs were posted in the work area and updated each month: one for overall effectiveness, and one for each indicator. These graphs allowed unit personnel to see changes in effectiveness over time. As one might imagine, both the feedback report and the graphs generated considerable interest when they were posted each month.

Once the feedback reports were circulated, a meeting was held with section supervisors and a representative from upper management to review the feedback report for the month. Areas of improvement were noted, and areas of decrease discussed. Reasons for the improvements or decreases were considered, and any longer-range trends were noted. This meeting also served as a basis for planning priorities for the next month, and for making changes to improve productivity.

This feedback of the information from the productivity measurement system was provided for five months in each of the five units. After this period, goal setting was added to the feedback. After five months of feedback plus goal setting, incentives were added in the form of time off for superior productivity. The details of these interventions are presented in a separate paper (Pritchard, et al., 1987).

Results

The results of this application of the productivity measurement system are presented in three sections: (1) results during development of the system, (2) effects on productivity, and (3) effects after the departure of the research team.

1. Results during development

One of the most interesting results that occurred during the development of the system was the change in the attitudes of the unit personnel. When we first started working with them, their attitudes toward the project were mixed. Although some unit personnel were positive, others were more skeptical. By the time system development was completed, however, almost all unit personnel had positive attitudes toward the effort. They were solidly behind the system, felt positive toward the researchers, and were quite disappointed when the start of feedback had to be delayed so that enough time had passed to establish a baseline.

Development of the system also resulted in a conscious examination of unit objectives, the development of possible measures of these objectives, and an evaluation of productivity expectations and limits. This process led unit supervisors to see numerous places where improvements could be made in the operation of the units. Naturally enough, they began to implement these changes. This created a real dilemma for the researchers. Although it was certainly worthwhile for the units to improve their effectiveness as a result of the development of the system, this improvement was occurring prior to the establishment of our baseline. If because of this, the baseline period showed higher effectiveness than it would have otherwise, this would decrease the size of any effect due to the productivity feedback. There was little the researchers could do about this dilemma. The units felt strongly that such changes should be made, and made them. They felt these changes were increasing their effectiveness, and this indeed seemed to be the case, based on what little data were available at the time. Interviews with supervisors indicated they believed that a substantial portion of this improvement was due to the process of developing the productivity measurement system. This suggests that the improvements in productivity that were evidenced in the interventions were, in fact, underestimates of the overall impact of the development and introduction of the feedback from the productivity measurement system.

Another issue concerning the development of the system had to do with the difficulty of the task for the personnel involved. We had expected the greatest difficulty to be in the development of the contingencies. We were quite mistaken. The development of the indicators was the most difficult.

One problem with the development of the indicators was that unit personnel tended to be very accepting of existing measures. This was not surprising, since they had not been trained to critically evaluate possible measures and were somewhat resistant to change. For example, one measure considered by Comm/Nav later turned out to contain elements due to supply and flightline maintenance. Only about 10% of the variance in the indicator was due to factors under the control of Comm/Nav. They had been using this as an indicator of Comm/Nav effectiveness, but when its actual content was identified, they dropped it as a measure.

Another problem in developing indicators was that unit personnel did not always see the implications of certain measures. For example, the Receiving section of MS&D must input information about each piece of incoming property into the computer. If a mistake is made in this process, a "reject" is later printed out by the computer. Unit personnel must then identify the cause of each of these rejects and correct the data. The indicator that was first proposed was the percent of rejects that were cleared from the reject list each day. At first, this seemed to be a quite reasonable index. However, after studying these lists, it became clear that some rejects were quite easy to clear, whereas others were extremely difficult. Thus, if percent cleared was used, the better the unit did on clearing rejects, the more they would be dealing with only the most difficult and time-consuming rejects. In the long run, this would automatically lead to a poorer percent cleared, and thus be a poor indicator.

In contrast, the development of the contingencies went very smoothly. As mentioned above, we had expected contingency generation to be quite difficult for the personnel involved. Contingencies are conceptually complex, and we felt the effectiveness scaling would be an especially difficult task for them to do. What in fact happened was that after the contingencies were explained, unit personnel had little difficulty in developing them. They seemed to grasp very quickly what contingencies were and how they would be useful. They told us that contingencies captured the way they thought about what they did, and in a more clear and comprehensive way than they had ever done before. They also said that the contingencies facilitated communication by providing a meaningful frame of reference for discussions of policy. Personnel became very involved in contingency development, and frequently made revisions in the developing contingencies between our visits to the base.

Reliability and validity were also assessed during the development of the system. Reliability was assessed by interjudge agreement on the contingencies. The Comm/Nav shop had two shifts in operation. Personnel from both shifts were involved with the development of products and indicators. To assess the reliability of the contingencies, we developed two independent sets of contingencies, one set with each shift. This produced two effectiveness scores for each value of each indicator: one set from the day shift and one set from the night shift. Correlations were calculated between the two sets of values for each contingency. For example, if a given contingency had eight levels of the indicator, there would be eight effectiveness scores developed by the day shift and an independent set of eight effectiveness scores developed by the night shift. The two sets of eight scores were then correlated, producing one correlation for each contingency. These correlations ranged from a low of .86 to a high of .99, with an average of .95. Thus, the reliability of the contingencies as measured by interjudge agreement was quite high.

The validity of the system was evaluated using five different productivity scenarios of hypothetical indicator data developed for Comm/Nav. This was done by selecting a reasonable value for each indicator in such a way that the different scenarios varied as to their overall effectiveness. Although the overall effectiveness of each of the five scenarios varied, the differences were not so large as to be completely obvious. Also, the changes in indicator values varied, but not always in the same direction. That is, although the overall effectiveness may have increased for a given scenario, some indicator values went down, while others went up. Six Comm/Nav supervisors were then given the indicator data on the five scenarios and asked to rank the scenarios as to their overall effectiveness. If the system is accurately reflecting relative importance, having supervisors rate the scenarios without knowing the scenario overall effectiveness scores should produce ratings which are highly correlated with overall effectiveness as calculated by the system.

These ratings were done approximately two months after the development of the system had been completed, but before any productivity feedback had started. Results showed a correlation of 1.0 between each supervisor's rankings and the overall effectiveness score calculated by the system. This constitutes additional evidence for the validity of the system.

2. <u>Effects on productivity</u>

Once the system was developed and a baseline established, the system was used to generate feedback. Next, goal setting was added to feedback, and finally, incentives were added. The overall effects across the five units are shown in Figure 9. As the figure indicates, overall effectiveness increased substantially over the baseline. The average increase over baseline was 50% for feedback, 75% for goal setting, and 76% for incentives. Figure 10 presents similar data for Comm/Nav alone. The data show average increases of 30% for feedback, 65% for goal setting, and 68% for incentives. Figure 11 presents the data for MS&D alone. Average percent increases were 54%, 77%, and 79%.

These data indicate a major increase in productivity. The effects were extremely large. In addition, the MS&D effects were consistent across sections. Each of the four sections showed a consistent pattern of increase for feedback, and even greater increases for goal setting and incentives. It would be worthwhile to present a sample of these changes in terms of indicator data to give a sense of their magnitude. The most important indicator for Comm/Nav, percent of repair demand met, had a mean of 88.5 during baseline. This mean became 90.5 during feedback, 93.9 during goal setting, and 92.6 during incentives. The most important indicators for MS&D were times to process Priority 2 property. The mean time to move a piece of property was 92.6 minutes during baseline. This figure was cut to 25.3 minutes during feedback, and to 19.6 and 19.8 minutes for goal setting and incentives, respectively.

These results were then examined in light of the level of manpower in the five experimental units during the period to determine if the changes in productivity that occurred during the treatments could possibly be due to changes in unit manpower. For Comm/Nav, the data examined were the total number of personnel in the shop each month. Our analysis indicated that mean number of personnel during baseline was 30.9. This figure increased slightly during feedback to 33.0, was 32.8 during goal setting, and dropped back to 31.0 during incentives. Since manpower levels during the period of highest gain in productivity were essentially equal to the level during baseline, we concluded that the increases in productivity were not caused by increased manpower. In MS&D, the data reviewed were total number of personnel, and the number of hours of overtime logged per month. Unlike Comm/Nav, MS&D routinely had considerable overtime. The mean level of manpower for MS&D was 51.8 for baseline, 53.7 for feedback, 48.4 for goal setting, and 49.2 for incentives. Thus, overall manpower decreased over the period of the treatments. In addition, number of hours of overtime decreased from a mean of 1,348 hours during baseline to 892 during

FIG. 9 PRODUCTIVITY OF ALL FIVE SECTIONS COMBINED

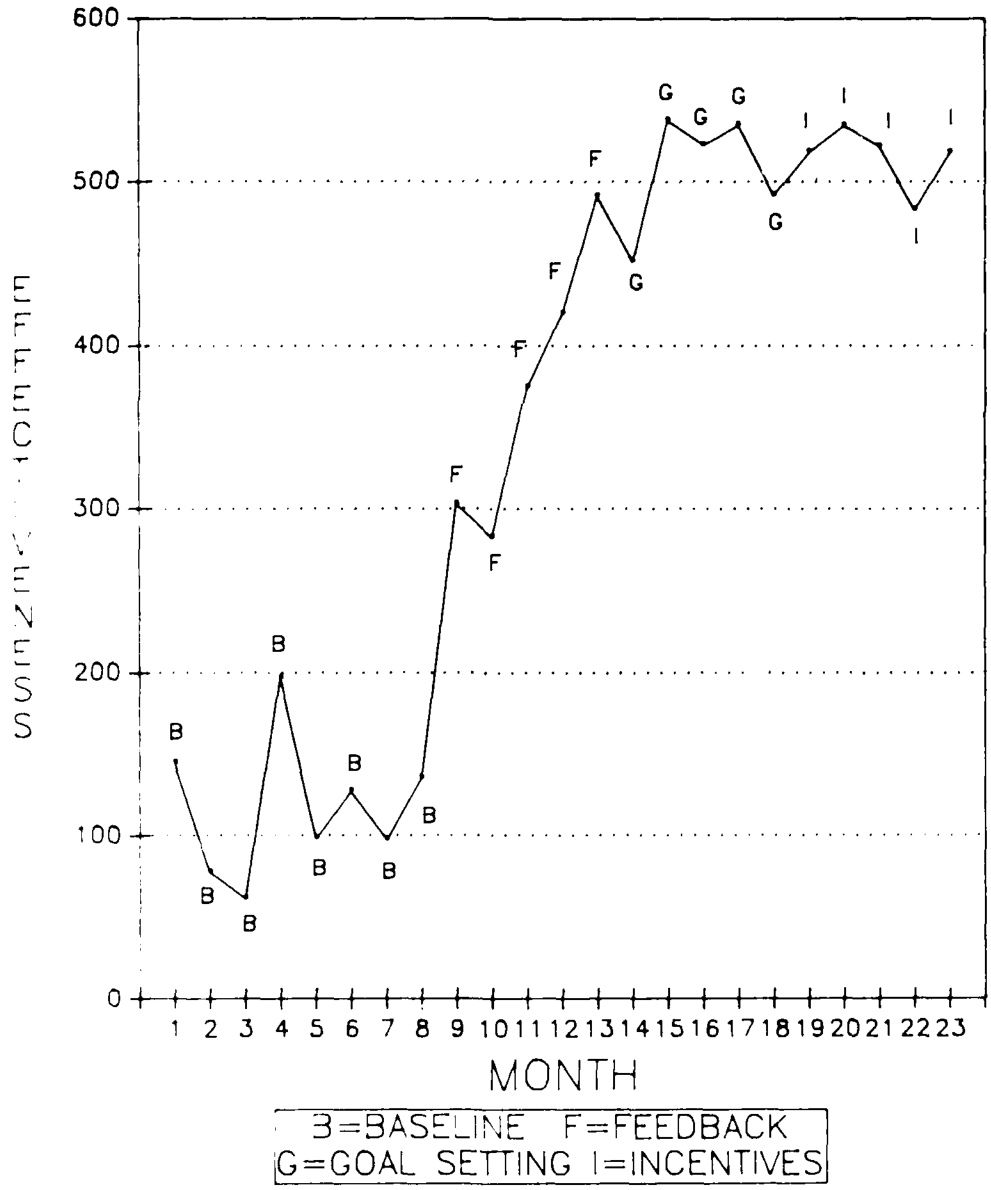
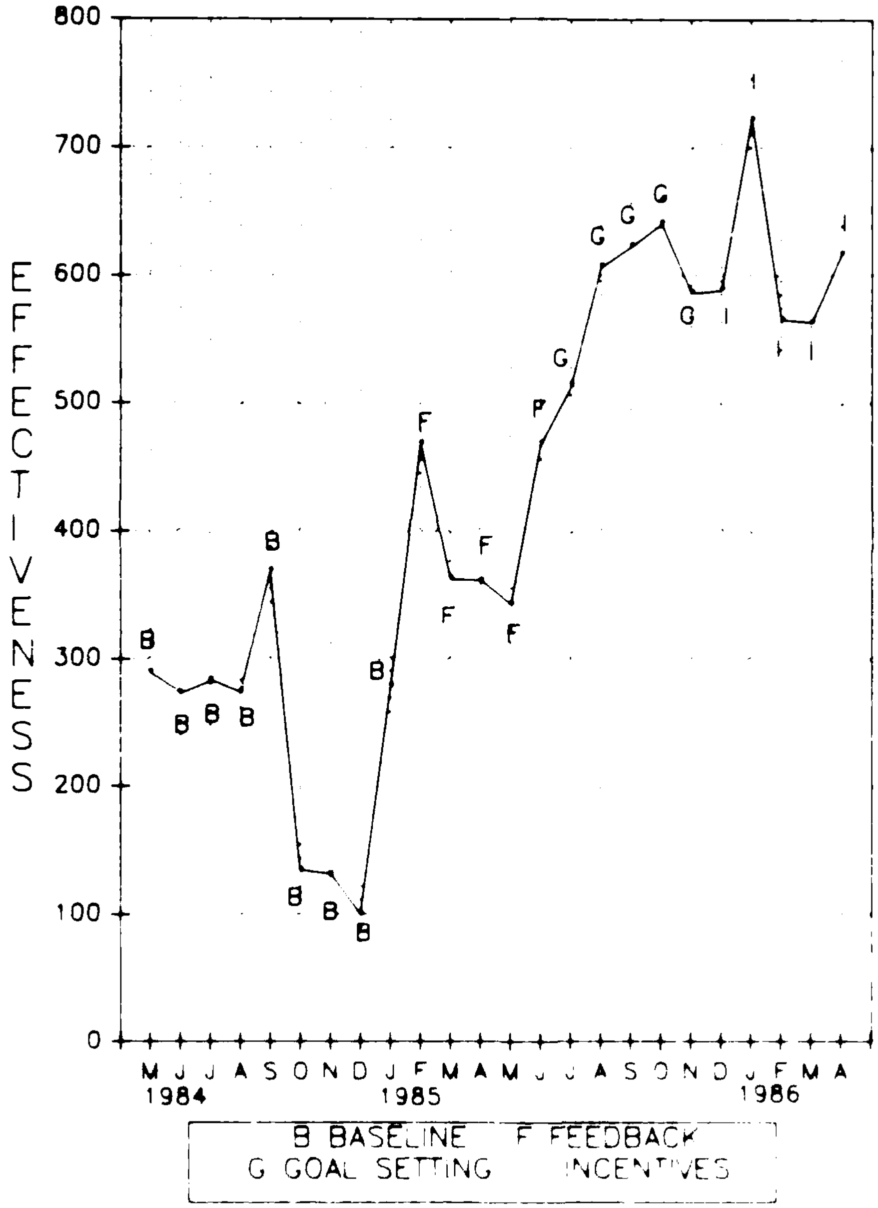
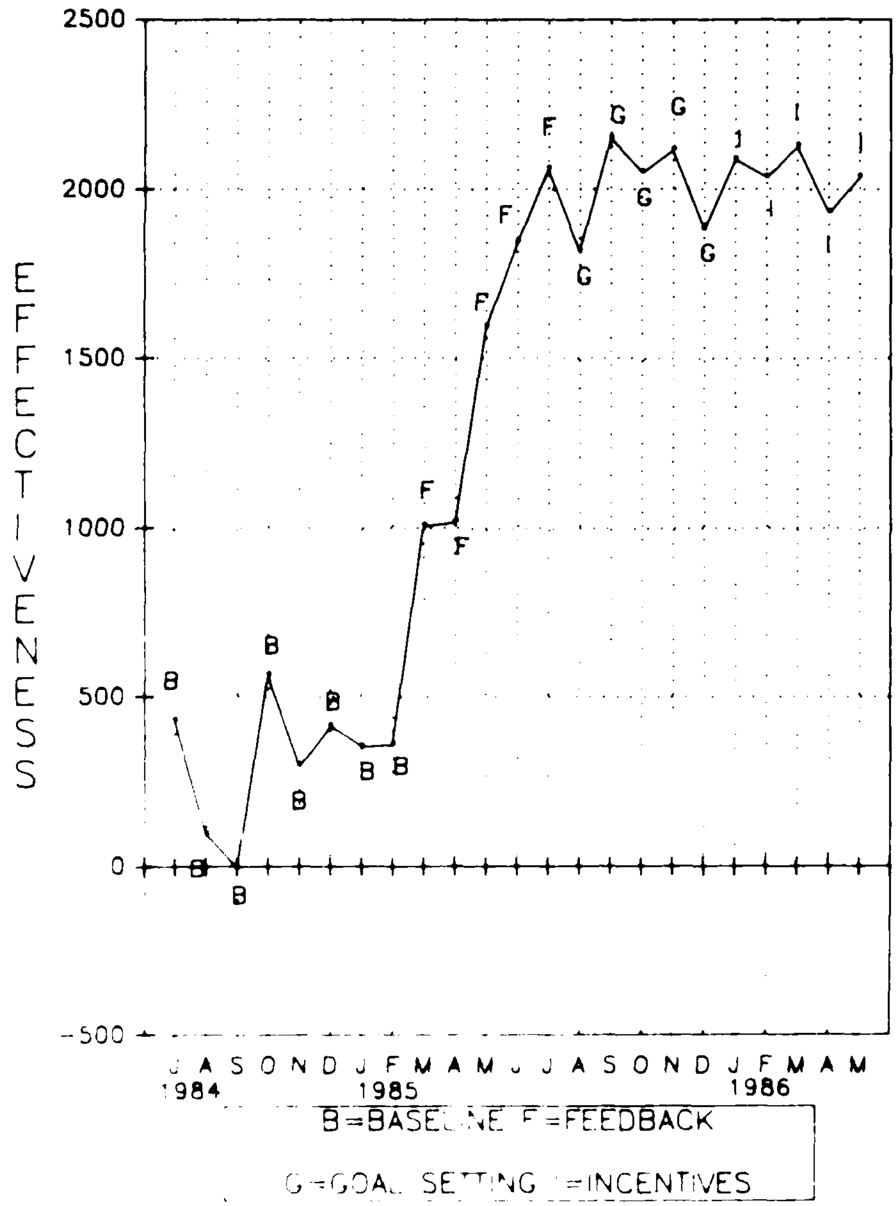


FIG.10 COMM/NAV PRODUCTIVITY



TOTAL SECTION EFFECTIVENESS OVER TIME

FIG.11 MS&D PRODUCTIVITY



TOTAL BRANCH EFFECTIVENESS OVER TIME

feedback, 404 during goal setting, and 416 during incentives. Thus, by the end of the treatments, overtime was less than one-third of what it was during baseline. These data indicate that the productivity gains that occurred were made with no increase in manpower in Comm/Nav, and a decrease in manpower in MS&D.

Results were also examined in light of productivity data from several sections similar to the experimental sections. These sections then, in essence, served as a control group, since they had no intervention from the research team. Control data for Comm/Nav consisted of 10 measures of productivity from eight maintenance units in the Component Repair Squadron. The data were collected from baseline and put into a composite measure to express overall productivity of the control groups. This composite measure was the sum of the ten measures. The mean value across the 10 measures during baseline was 317; it dropped to 295 during feedback, and rose to 377 and 365 during goal setting and incentives, respectively.

These results show that the productivity of the control sections decreased somewhat during the Comm/Nav feedback period, and increased thereafter. This would suggest that the productivity increase during feedback in Comm/Nav was not due to wider organizational changes since other squadron units actually decreased during this period. Furthermore, the increases during the Comm/Nav goal setting and incentives periods were not present across all units. These increases were brought about primarily by large increases in productivity on two of the ten control measures.

Productivity data on four control measures were examined for MS&D. These measures reflected overall functioning of the Supply Squadron (exclusive of MS&D), and one other unit outside of Supply. The mean of these four measures during baseline was 516; it was 512 for feedback, 511 for goal setting, and 518 for incentives. Thus, there were essentially no changes in productivity for the sections outside of MS&D.

Taken together, the control data indicate that the effects on productivity that occurred in the experimental units cannot be explained by wider organizational changes in productivity.

Data were also collected on subjective reactions to the system. All incumbents and supervisors (N=97) were administered a survey after several months of experience with the feedback system. They were asked to rate different aspects of the feedback system using 5-point Likert scales with response formats ranging from Strongly Agree to Strongly Disagree. Responses to survey items were uniformly positive. Table 1 presents the items and the percentage of

Table 1. Subjective Evaluation of the System

ITEM	PERCENT AGREE OR STRONGLY AGREE	PERCENT DISAGREE OR STRONGLY DISAGREE
1. The feedback system tells me how good a job I am doing.	64	4
2. The feedback system tells me how good a job the section is doing.	87	1
3. The feedback system helps me see the section's priorities.	77	6
4. The feedback system helps the section be more productive.	61	4
5. A system like this would help other Air Force bases be more productive.	62	4
6. The feedback system is clear and understandable.	58	7
7. It was worth the effort to develop the feedback system.	64	10
8. It was worth the effort to keep the feedback system in operation.	62	10
9. The information about section performance that goes into the feedback system is accurate.	52	13
10. The feedback system gives a good measure of productivity.	64	13
11. Overall, I like the feedback system.	62	9
12. I would prefer not to have this feedback system at the next organization I work in.	13	54
13. The feedback system is a better way of measuring productivity than what the section used to have.	75	0
MEANS =	64.7	7.2

respondents who agreed or strongly agreed and the percentage who disagreed or strongly disagreed. After reverse scoring the negatively worded item (#12), the mean percentage of respondents across all items who Strongly Agreed or Agreed was 64.7%, while the corresponding percentage who Disagreed or Strongly Disagreed was 7.2%. Clearly the response to the system by those who used it was very positive.

3. Effects after the departure of the research team.

To further evaluate the system, we also looked at what happened to the system after the departure of the researchers. Once the 5-month incentive treatment was over, our on-base responsibilities officially ended. While we were on base for a variety of purposes after this time, the units had no commitments to continue the system. However, a significant indication of the system's value is that each of the five units elected to continue the system after the researchers left. This meant that they had to commit their own resources to put together the indicator data and run the programs producing the feedback reports. In addition, we were asked by both Comm/Nav and MS&D management to develop the system in other units in their respective squadrons. Although we did not have the resources to do so, it did indicate the value that the units placed on the system.

At the end of the formal incentives treatment, units were asked if they wished to modify the system. If they did want to, we agreed to assist if it became necessary. Comm/Nav and two sections of MS&D elected to make changes. In all three cases, the changes were to eliminate some indicators from the system. The indicators that were removed were those with very flat contingencies (indicating they were not very important), those pertaining to activities that the unit was no longer going to perform, or those that appeared to be under such good control that they were no longer important to measure.

A major strength of the system is that it can accommodate changes readily. As changes occur in policies, procedures, or resources, changes will need to be made in the system. This can be done by eliminating indicators, redefining them, or altering the scaling of contingencies. Thus, the system can easily be altered to changing conditions. However, it must be understood that after such changes are made, the new effectiveness scores are no longer comparable to the old scores. For example, if indicators are dropped, the same actual productivity will show up as lower overall effectiveness since some effectiveness points are lost to the deleted indicators. This makes the interpretation of effects over time difficult until a new "baseline" is established. A new baseline is established by taking the newly revised system of indicators and contingencies and calculating its overall effectiveness for several months prior to the revision. For example, if some indicators are dropped, it is a straightforward matter to go back to the indicator data from prior months and calculate what overall effectiveness would have been if those indicators had not been included. This then becomes the new baseline, and the effectiveness scores after the revision are completely interpretable. If indicators are added, it is a matter of recalculating what the overall effectiveness for prior months would have been had these indicators been included. This is a simple matter, provided historical data are available.

One feature of the system that needed to be explored was whether the units would be able to use the system after the researchers left. As part of that process, the units needed to be able to make changes in the system, since the need for changes would undoubtedly occur in the future. Thus, we were particularly interested in their ability to do this. In making these changes, it became clear that the management of the units understood the system fairly well, and with help from the research team were able to make the changes. With this help, they were able to eliminate the indicators they wished, adjust the calculations of effectiveness to take the removal of the indicators into account, and recalculate past effectiveness to establish a new baseline. Our assessment is that Comm/Nav could now make such changes completely on their own, and MS&D could with minimal help.

Before we left the base, the units wanted us to train them to use the system, so that they could continue using it after we left. This proved to be a fairly simple task in Comm/Nav. By the end of the incentives treatment, they had already taken over the collection of all the data that went into the system. They had only to be trained on using the microcomputer programs that were used to calculate effectiveness scores and generate the feedback reports. This was done readily, and other than an occasional question, they have been operating the system completely on their own for several months.

The training in MS&D was more involved. Although their intent to take over the system had been frequently expressed for some time, the actual commitment of personnel was not made until the end of the incentive treatment. Thus, the training could not be done gradually over many months, as would have been optimal. In addition, the task of preparing the feedback reports in MS&D was more difficult than in Comm/Nav. The MS&D reports required the entering of data showing the amount of time it has taken to move property in the warehouse. Someone must take several hours each month to enter these data and run the program that calculates the mean times for the indicator data. During the interventions, this was done by the research team. After our departure, it had to be done by MS&D personnel. Thus, they had to learn how to enter these data and run the program, and also learn how to generate the feedback reports. Because of these factors, it took longer to train MS&D to take over the system, but they did eventually learn to do so successfully.

A final consideration here is what happened to productivity after the units took over the system themselves. The results are shown in Figure 12. This figure not only indicates the effects after the units took over system operation, but also demonstrates the capability of the system to generate a new baseline when changes in the system are made. Since both Comm/Nav and some sections of MS&D had deleted indicators from the system, we had only to recalculate the overall effectiveness scores back in time in order to develop a baseline or comparison. In this case, all changes in the system were made the month following the incentive treatments, when our involvement in the interventions ended. To develop a baseline, we recalculated the overall effectiveness data for the five months of incentive treatment. This *adjusted* score is the overall effectiveness score that the units would have had during the incentive treatments if they had been under the revised system.

Based on the revised system, the mean overall effectiveness score under incentives for Comm/Nav was 519, and the mean after they took over the system was 556. For MS&D, the mean under incentives was 1857, and the mean after they took over was 1792. For both units combined, the respective means were 2376 and 2348 during and after incentives, respectively. Thus, the data indicate that productivity remained approximately as high after base personnel had responsibility for the system as it had been when the system was operated by the research team.

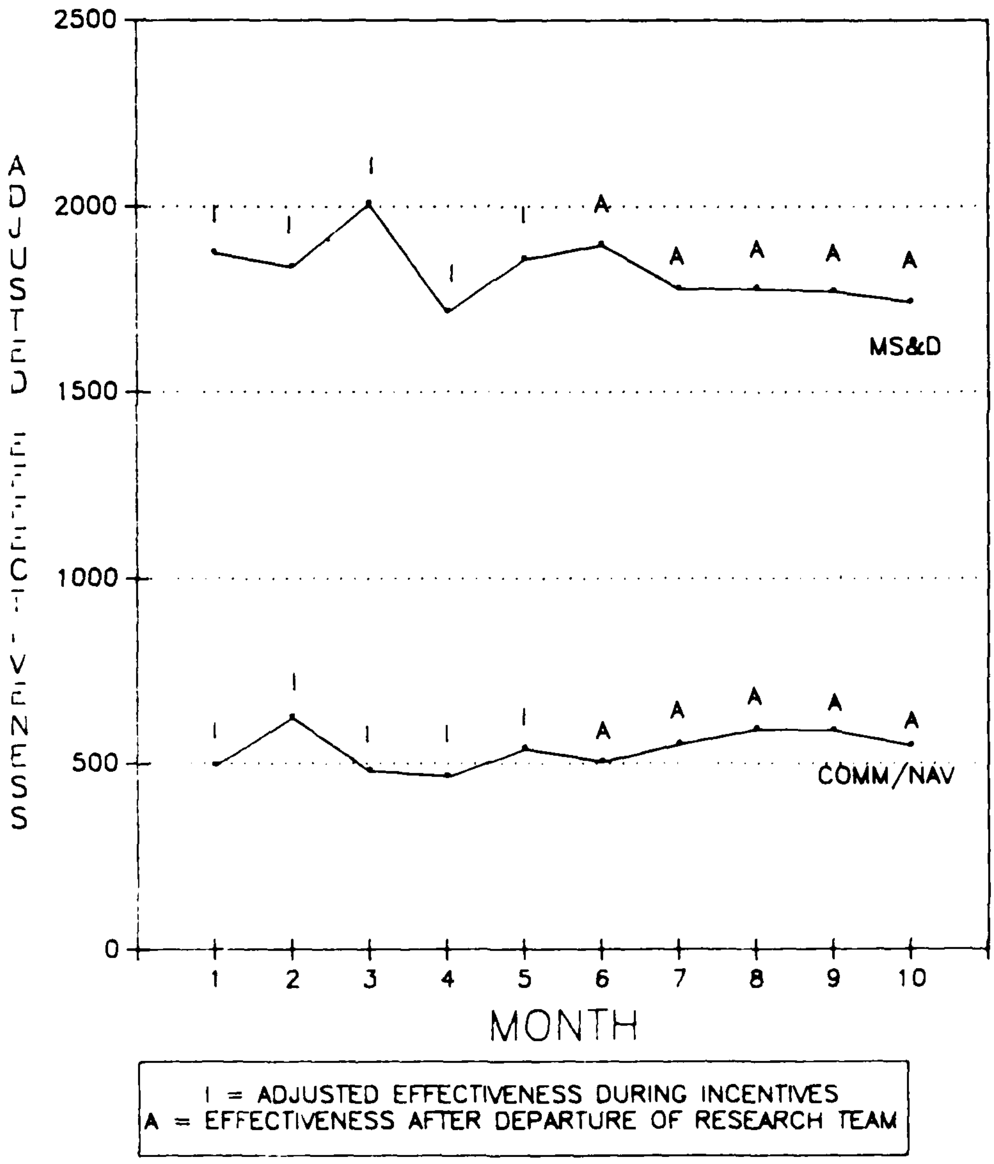
V. DISCUSSION

It seems clear that the productivity measurement system and the resulting feedback system were extremely successful. We believe this success can be attributed to the following factors: structure of the system, motivational value, organizational accountability, and feedback.

Structure Of The System

One of the strengths of the system is that it allows productivity to be expressed as a single index. Such an index is useful to management, supervision, and incumbents since it presents information in a highly convenient form. One group of supervisors told us that even if all the data on the indicators had been previously available, these data would be in different places and hard to put

FIG. 12 PRODUCTIVITY AFTER THE DEPARTURE OF THE RESEARCH TEAM



together, and it would have been impossible to derive an overall sense of how the unit was doing. This system combines all the data in an easy-to-understand manner.

The single index also has other advantages. It can be used as a basis for the implementation of interventions. For example, one of the problems with goal setting programs is in their application to complex jobs. The dilemma is whether to have goals set in all the important areas of the job and thereby have a very complex and hard-to-understand goal program, or have goals set in only one or two important areas and run the risk of having the other important activities somewhat ignored (Duerr, 1974). With a single index of productivity, the problem is solved because it includes all the important functions; only the overall index need be used in setting goals. The single overall index makes many other interventions, such as incentives and gainsharing, much easier to implement as well.

Another advantage of the single index is that it facilitates tracking productivity over time. In addition to the obvious motivational implications of this feature, it also allows the unit and its management to evaluate the effects of any intervention on the unit. Interventions could range from a change in management, in workflow, or in the type of equipment used, to interventions such as goal setting or incentives.

A productivity measurement system should also have subindices of productivity to provide information on the components of overall productivity. This information has value both as positive feedback and as information for productivity improvement.

Another important aspect of the system was that acceptance by unit personnel was high, at least by the end of its development. Their acceptance was likely due to the fact that the system was tailored to their particular situation, and they had a major hand in its development. They had a sense of ownership which likely decreased any tendency toward resistance to implementation of the system (Hurst, 1980; Tuttle & Sink, 1984). Acceptance was even further enhanced once the system started generating feedback for the units. Personnel found the information useful in doing their work.

The system was also successful because of its validity. First of all, the system was valid in that it was complete; that is, it included all the important aspects of the units' work. Its completeness was achieved through the careful review and approval process used. If data that had not been collected in the past were needed to make the system complete, the new data were collected. A major index of the system's completeness was that, across the five units, not one indicator needed to be added during the 20 months of experience with the system.

The system is also valid in terms of its ability to accurately reflect how well the unit is functioning. For the system to be valid, the products must be correct, the indicators must be good indices of the products, and the contingencies must be accurate. The structure of the system allows for high validity because the relative importance of different activities is maintained, and non-linearities are preserved. The iterative process of development of products, indicators, and contingencies promoted the type of repeated review that maximizes accuracy. Higher management's approval ensured that the system was an accurate reflection of organizational policy. Finally, the reliability and validity data obtained were highly supportive of the validity of the system.

Another apparent reason for the success of the system was that it allowed for direct comparison across units. That is, the system allows units that have totally different functions to be directly compared as to their level of productivity. This feature was utilized considerably by the management of MS&D. Having each unit's overall effectiveness expressed in terms of percent of maximum possible effectiveness, it was easy for management to assess how well the units were doing relative to each other; for example, to make decisions about resource allocation across units.

Another important feature of this system is its flexibility. The system is flexible in that it can accommodate both effectiveness and efficiency approaches to productivity. The effectiveness approach is the one used here, and the effectiveness scores are an expression of output relative to expectations. However, the system can also accommodate an efficiency approach by incorporating efficiency into the indicators. For example, a measure of Comm/Nav monthly repair demand met divided by the manhours available that month would be an efficiency measure of labor productivity. Thus, efficiency can be included in the indicators.

In addition, efficiency can be included by taking the overall effectiveness measure and dividing by total inputs. This becomes a measure which combines efficiency with effectiveness, and may be a very valuable approach in many situations. For example, overall effectiveness could be divided by the total costs over which the unit has control (costs of manpower, supplies, etc.) to obtain a measure of cost effectiveness. Unit personnel would then be expected to maximize cost effectiveness by increasing effectiveness, decreasing costs, or both. Yet another approach would be to divide overall effectiveness by the total costs of operating the unit whether controllable by the unit or not. Such a cost effectiveness measure would be a major component in a management information system. Finally, multiple measures could be used to get a more complete picture of the functioning of a unit.

The system is also flexible in terms of its ability to accommodate changes within an organization. There will be changes over time in any organization's objectives or procedures. These could be brought about by changes in policy, changes in technology, or refinements in operations. The productivity measurement system must be able to accommodate these changes when they occur. This system can accommodate changes through modification of any of its components. Products, indicators, or contingencies can be changed. For example, if a unit decided to change from an emphasis on quantity to an emphasis on quality, the contingencies for the indicators of quality could be made steeper, and the slopes of the quantity indicators less steep. Or, if expectations change, changes can be made in the zero points of some of the contingencies, or the entire contingency can be rescaled.

An important advantage of the system is that it can be applied to any level of the organization, and can be aggregated to larger and larger parts of the organization, as we did in MS&D. In principle, this aggregation could be continued until the entire organization is included, and its effectiveness could be expressed in a single number.

In our field test, we chose to develop the system at the lowest level in the organization. We did so because we wanted to affect the motivation of unit personnel and, from our previous work on feedback, had concluded that the maximum impact on motivation comes about when feedback is directly relevant to the specific tasks that are being done. It is also possible, however, to develop the system at much higher levels of the organization. One could develop products, indicators, and contingencies at the squadron level, for example. Such an approach would not have the motivational features of being able to give specific feedback down to the lower levels, but it would serve as a basis for understanding and monitoring the productivity of the larger unit.

The ideal strategy would be to start at the lowest level and build the system for each section, then aggregate up to the level of the higher unit (e.g. the squadron or the wing). This would have the advantage of providing motivational impact to the individual sections and branches, and also provide information on the functioning of the entire squadron or wing.

Motivational Value

A second factor that seemed to make the system effective was its strong motivational value. First of all, the system resulted in a <u>clarification of roles</u>. The process of developing, refining, and getting management approval for the products, indicators, and contingencies forced role clarification. When the process was finished, the units had a clear picture of what their objectives were, what they should be focusing on to achieve these objectives, what was expected of them in each area, and what good and bad productivity were in each area. Thus, they had a far better understanding of their roles than they had before the development process started. This role clarification process likely had positive motivational effects in and of itself.

The use of the system also dramatically increased the amount and quality of <u>feedback</u> unit personnel received. With the system, they received more accurate feedback, and more positive feedback; when the feedback was negative, it was provided in a useful, non-personal form. The positive feedback seemed particularly important. Personnel indicated that the system represented one of the few times they had been told they were doing a good job. The productivity measurement system provides a great deal of positive feedback. When the unit is doing well or when it is improving, the reports show this. This positive feedback is a very important feature of the system.

Another motivational feature of the system is that personnel <u>could see the results of their efforts</u>. Most jobs are structured such that doing a better job does not show up in any measurable way. This can become very demotivating. The frequent feedback provided by the system seemed to improve the connection between individual effort and unit productivity. Personnel could see the effects of their efforts to improve productivity.

The system also allowed for <u>competition across units</u>. Using this system, units can be compared in terms of percent of maximum effectiveness, and can thus compete on this basis. In MS&D, competition was clearly present between sections, was friendly in nature, and seemed to have a positive effect on productivity.

One of the most important features of the system was that it allowed personnel to <u>focus on the same objectives</u>. Before implementation of the system, different supervisors and different levels of management focused their efforts on different things. Furthermore, what was high priority changed frequently. One week a great deal of effort would be put on one thing, and the next week something different would be top priority. This is a problem common to most organizations. It is brought about by people having different ideas about the units' priorities. It is especially problematic when there is no agreement on priorities by different levels of management. Constantly changing priorities create serious problems for unit personnel.

The process of developing the productivity measurement system seemed to reduce this serious problem dramatically. In essence, all levels of management had agreed on what was and was not important. All levels had an opportunity to see the perspective of the others, all had agreed on what was important, and a concerted effort could then be made to accomplish the organization's objectives.

Organizational Accountability

Another factor in the system's effectiveness was that the system made units accountable for their productivity. Units can be held more accountable when their productivity is measured and reported on a regular basis. Concrete performance data are hard to ignore. The fact that the data exist and will continue to be provided generates a source of motivation for unit personnel to want to look good; also, they know that they will have to answer for it if they do poorly. However, whenever productivity data are presented objectively about problem areas, as they are in this system, there is much less of a tendency to make excuses, and more of a desire to try to find positive solutions.

Another aspect of accountability is that the system allows for a way of assessing the effectiveness of supervisors and managers. Their major responsibility is to effectively use the resources under their control to maximize the achievement of the organization's objectives. Thus, one very useful index of their effectiveness is the effectiveness of their unit.

Feedback

Another apparent factor in the effectiveness of the system was that it provided considerable information for identification and diagnosis of problems. The feedback reports showed if productivity had started to slip in a given area. This allowed the unit to deal with the problem before it became serious. Prior to receiving feedback, a problem could become much more pronounced before it came to the attention of the unit. In addition, the feedback reports aided in diagnosing the causes of problems. Finally, as mentioned above, the feedback allowed the unit to know when a problem was fixed.

The system also gave information on appropriate priorities. The section of the feedback report that indicates where the maximum increases in effectiveness would result served as a guide for setting priorities and allocating resources for the upcoming month.

VI. CONCLUSIONS

A number of conclusions can be drawn from this research. First we will discuss specific research conclusions concerning the capabilities of the system. We will then discuss issues concerning the productivity measurement system development process. Finally, we will review other potential applications of this technique.

Specific Research Conclusions on System Capabilities

The system evaluated in the present effort appears to be a very effective way of measuring productivity. Its implementation is quite feasible, unit personnel were cooperative in developing the system, and it showed good psychometric properties. In addition, the system development process itself seemed to have a [illegible] effect on unit functioning even before feedback was instituted.

The feedback that resulted from the productivity measurement system had [illegible] strong effect on productivity. An average gain in productivity of 50% occurred across the five units during feedback, and gains of 75% and 76% occurred [illegible] feedback with goal setting and incentive interventions, respectively. [illegible] effects of the system have lasted over time. Specifically, the positive productivity results continued for the 15 months that the research team was [illegible], and have continued for at least 5 months after that.

[illegible] system allowed for aggregation across units so that an integrated [illegible] be developed across the four sections making up a branch. This [illegible] quite simple once the basic system is developed in each section. [illegible] application of this aggregation to much larger and more complex organizations seems quite feasible.

[illegible] units using the system evaluated it very positively. Expressed [illegible] toward the system ranged from positive to very positive. The members [illegible] were quite proud of the system and their productivity improvements. [illegible] unit members frequently showed the system to people visiting the [illegible], after the experimental period, all of the units continued using [illegible] their own, and management wanted to expand it to other units as

The development of the system had positive effects for units that were quite different from one another in terms of the nature of the work. The activities were quite different between Comm/Nav and MS&D, as well as among the four sections of MS&D. There were also great differences in the type of organizational structure, and the work flow. The units varied considerably in size, and the personnel varied considerably as to their academic and technical training background. They also differed as to their initial levels of performance. Yet, with all these differences, the system was developed and worked extremely well in each unit. This adds considerable support to the generalizability of the approach.

The system also offers a number of other benefits. It allows for the direct comparison of the productivity of different units to each other. It can be used with both effectiveness and efficiency approaches to productivity. In addition, it can be applied to any level of the organization, allows for competition between different units, helps identify the priorities for increasing productivity, and serves as an excellent basis for evaluating any changes made in the organization.

Conclusions About the Productivity System Development Process

In doing this research, a number of conclusions were drawn concerning the process used for developing productivity measurement and enhancement programs in Air Force environments.

One is the importance of having the personnel who are going to be using the system be heavily involved in its development. The perception of unit personnel was that some previous programs imposed from above had not been effective because these programs were not designed with an appreciation for their unique needs and environment. It is much more effective to have heavy involvement from the personnel that are going to be using the system, so that the final product will fit their needs and they do not feel that it is yet one more project imposed from above.

It also seems more effective to develop such programs from the bottom of the organization up. The lower levels of supervision know the most about the functioning of the unit, and what are the real critical issues. In addition, these are the people that will make the system work. It is important to have their involvement and knowledge from the start. It is also important to have higher-level involvement to approve the system. This should be done as the system is being developed, not delayed until the system is complete. We believe our technique of getting approval on products and indicators before starting contingency development proved very valuable. It not only helped clarify policy earlier in the process, but also helped prepare everyone for the eventual implementation. This approach gave all levels of the organization a chance to learn about the system as it was being developed so that they would know how to use it when it was finished. It also served to generate considerable eagerness at all levels to receive the first feedback from the system.

It also proved very effective to have unit personnel who developed the system defend it when it was presented to higher management. They were much more knowledgeable than the researchers about the subtleties of unit operation, and could better address management's questions and concerns. Also, it gave them a chance to present their perceptions of optimal policy. Finally, their sense of ownership of the system was increased by their defense of it.

We also learned the importance of using a multiple meeting structure that allows for an iterative approach to the development of the system. Our strategy was to summarize the results of each meeting and present them at the next meeting to assure consensus was reached. In this way, unit personnel had time to think about the issues, discuss them among themselves, and be prepared to approach the question with fresh perspectives at the next meeting. Personnel need time to adjust to the idea of completely capturing what they do in a productivity measurement system. They need to think about it, consider its implications, and be able to discuss it among themselves thoroughly. We believe that a quality system could not have been built without multiple meetings separated in time.

Care must be used during development to ensure that the resulting indicators are measures which are under the control of the unit (Harst, 19[illegible]). The researchers needed to frequently remind unit personnel to assess whether they had control over a given indicator. Unless the unit has full control over the work being performed, including measures of performance of that work would be counterproductive in that they would decrease the motivational [illegible] of the system.

Finally, it is important that the researchers [illegible] with the operational personnel. Some personnel were [illegible] and had questions about our credibility. [illegible] who were not experts in [illegible] at best resulted in extra work and at worst had a negative [illegible] effectiveness. Taking the time to really [illegible] what they [illegible] know them, went a long way toward increasing [illegible]

Other Applications Of This Approach

The productivity measurement system used in this research has a number of possible applications. We have discussed how the system can be aggregated to larger and larger units. In theory, the system could be developed for all units of a very large organization. There would be an overall effectiveness index for the entire organization, and for each subunit down to the lowest level in the organization. It would be theoretically possible, for example, to develop such a system for the entire Air Force. Although such a project would take a tremendous amount of effort, it is feasible.

Another possible application is in the area of management information systems (MIS). An MIS is designed to provide information to high-level decision makers for planning and resource allocation purposes. The problem with the information in an MIS is that there is frequently too much of it. It is hard to separate out the crucial from the not so crucial. It is particularly hard to combine the information in a way that will facilitate decision making. One application of our approach would be to combine the individual pieces of information in the MIS by using the contingency approach. For example, if there were ten pieces of information about the functioning of a given unit, developing contingencies relating these ten measures to overall functioning would make the interpretation of how well that part of the organization was doing much easier. It would also make comparisons of that unit to other units easier since each would be measured on the same metric.

Another possible application of this approach is in the area of criterion development. The multiple criteria issue has been a problem for years (Dunnette, [illegible]; Schmidt & Kaplan, 1971). The issue is that for any job there are multiple activities and multiple indices of performance. There may be ratings on different dimensions, job samples, and measures of output. The problem is how to combine these different indices into a single measure of performance.

[illegible]

[illegible]

developed for each type of job. This preserves the relative importance of the different criteria and also preserves non-linearities. A single index of performance could then be developed for use as a criterion.

Another aspect of criterion development that deserves special attention is criteria of managerial performance. Developing good managerial criteria has always been a difficult task. Our system could be of use here. We assume that the most important task of the manager is to use the resources under his/her control to meet the objectives of the unit in the most effective manner possible. This implies that one of the most important aspects of the performance of the manager is the productivity of the units under his/her control. Thus, developing the productivity measurement system for a set of units in the organization also creates a criterion for the performance of the units' managers.

The approach can also be used as the basis for developing performance appraisal systems. As with the criterion in selection, one of the problems in performance appraisal has been the lack of an overall index of performance. Typically ratings of performance have been made on performance dimensions, and an overall evaluation made as a separate global rating, or as some summation of the ratings of the individual dimensions.

We would apply the logic of the productivity measurement system to develop the appraisal system. Once the dimensions were identified, anchors for each scale would be developed that were analogous to indicators. That is, they would represent behavioral indicators of how well the person was doing on that dimension. These would then be used to develop contingencies. Once this was done, the overall performance rating would then be determined by summing the effectiveness scores for each dimension in a fashion similar to the way in which overall effectiveness is determined in the productivity measurement system.

This would have several advantages. First, an overall measure of performance would be generated. It would weight the dimensions of performance according to their importance, and preserve non-linearities. Second, by keeping the ratings to the more molecular dimension ratings, the rater would not have to make the more molar judgment about overall performance. This overall judgment would be generated mechanically. This could have the effect of decreasing rater errors. Next, the very process of developing the system should help in role clarification, as it does when used as a productivity measurement system. Finally, such a system would be useful for performance feedback. The system itself would communicate what is important and what is less important, and what level of performance is expected in each area. The ratings from the system would indicate what the person did well on, and not so well on. It would give an overall index of performance. The priorities that come out of the system for increasing performance would be a good source of information for the ratee, and a basis for performance counselling by the supervisor.

Finally, there is a set of applications for this approach that is more general in nature. The approach of indicators and contingencies can be used whenever there are multiple pieces of information that must be combined into an overall index. For example, promotion in an organization could be based on such things as performance ratings, experience, training, and test data. This set of very divergent information must be combined into a single evaluation of promotability. The system used here is an ideal way to combine this information, and has the added advantage of producing a clear specification of promotion policy. There are many situations where this combination of information takes place. Examples include evaluating alternatives in decision making, medical or psychological diagnosis, accepting candidates for training programs, and making lending decisions.

In conclusion, this approach to measuring productivity appears to be a very good one. It is feasible and effective, it enhances productivity when used as feedback, and it has a number of other desirable features. It can be applied to much larger organizational units. Finally, the basic logic has a number of applications beyond strict productivity measurement.

REFERENCES

Alluisi, E.A., & Meigs, D.K., Jr. (1983). Potentials for productivity enhancement from psychological research and development. American Psychologist, 38, 487-493.

American Productivity Center (1981). Productivity perspectives. Houston, TX: Author.

Balk, W.L. (1975, March-April). Technological trends in productivity measurement. Public Personnel Management, 128-133.

Campbell, J.P. (1977). On the nature of organizational effectiveness. In P.S. Goodman, J.M. Pennings & Associates (Eds.) New perspectives on organizational effectiveness (pp. 13-55). San Francisco: Jossey-Bass.

Campbell, J.P., Bownas, D.A., Peterson, N.G., & Dunnette, M.D. (1974). The measurement of organizational effectiveness: A review of relevant research and opinion (Contract N00022-73-C-0023). Minneapolis, MN: Navy Personnel Research and Development Center.

Coulter, P.B. (1979). Organizational effectiveness in the public sector: The example of municipal fire protection. Administrative Science Quarterly, 24(1), 65-81.

Duerr, E.C. (1974). The effect of misdirected incentives on employee behavior. Personnel Journal, 53(12), 890-893.

Dunnette, M.D. (1963). A note on the criterion. Journal of Applied Psychology, 47, 251-254.

Engel, J.E. (1977). A study of the relationship between worker attitudes and organizational effectiveness in an air logistics center maintenance directorate (AFIT-LS-3-77, AD A044586). Wright Patterson AFB, OH: Air Force Institute of Technology.

Felix, G.H., & Riggs, J.L. (1983). Productivity measurement by objectives. In W.B. Werther, Jr., W.A. Ruch, & L. McClure (Eds.), Productivity through people (pp. 349-356). St. Paul, MN: West.

Goodman, P., & Pennings, J. (1976). Toward a framework of organizational effectiveness (AD-A029 327). Pittsburgh, PA: Carnegie Mellon University.

Hanes, L.J., & Kriebel, C.H. (Eds.) (1978). Research on productivity measurement system for administrative services: Computing and information services (Vol. 2). (National Science Foundation Grant No. APR75-20546).

Hurst, E.G. (1980). Attributes of performance measures. *Public Productivity Review*, 4(1), 43-50.

Joint Financial Management Improvement Program (1976). *Productivity programs in the Federal government, Vol. I: Productivity trends and current efforts*. Washington, DC: Author.

Katzell, R.A., Yankelovich, D., Fein, M., Ornati, D.A., & Nash, A. (1975). *Work productivity, and job satisfaction*. New York: The Psychological Corporation.

Keeley, M. (1978). A social-justice approach to organizational evaluation. *Administrative Science Quarterly*, 23(2), 272-292.

Kim, J.E. (1980). Cost-effectiveness/benefit analysis of post-secondary occupational programs: A conceptual framework. *Planning and Changing*, 11(3), 150-165.

Mahoney, T.A., & Frost, P.J. (1974). The role of technology in models of organizational effectiveness. *Organizational Behavior and Human Performance*, 11, 122-138.

Mahoney, T.A., & Weitzel, W. (1969). Managerial models of organizational effectiveness. *Administrative Science Quarterly*, 14, 357-365.

Mendelow, A.L. (1983). Setting corporate goals and measuring organizational effectiveness - A practical approach. *Long Range Planning*, 16(1), 70-76.

National Center for Productivity and Quality of Working Life (1978). *Total performance management: Some pointers for action* (NTIS No. PB300249). Washington, DC.

Naylor, J.C., Pritchard, R.D., & Ilgen, D.R. (1980). *A theory of behavior in organizations*. New York: Academic Press.

Norman, R.G., & Bahiri, S. (1972). *Productivity measurement and incentives*. London: Butterworths.

Peeples, D.E. (1978). Measure for productivity. *Datamation*, 24(5), 222-230.

Perrow, C. (1961). The analysis of goals in complex organizations. *American Sociological Review*, 26, 854-866.

Price, J.L. (1968). *Organizational effectiveness*. Homewood, IL: Irwin.

Pritchard, R.D., Jones, S.D., Roth, P.L., Stuebing, K.K., & Ekeberg, S.E. (1987). *A field test of effects of feedback, goal setting, and incentives on organizational productivity* (AFHRL-TR-87-3). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Schmidt, F.L., & Kaplan, L.B. (1971). Composite vs. multiple criteria: A review and resolution of the controversy. *Personnel Psychology*, 24, 419-439.

Seashore, B.E., & Yuchtman, E. (1967). Factorial analysis of organizational performance. *Administrative Science Quarterly*, 12, 377-395.

Sibson, R.E. (1976). *Increasing employee productivity*. New York: AMACON.

Steers, R., Porter, L., Mowday, R., & Stone, E. (1975). *Problems in the measurement of organizational effectiveness* (TR No. ' Contract No. N00014-76-C-0164, AD A018 709). Eugene, OR: University of Oregon, Office of Naval Research.

Tuttle, T.C. (1981). *Productivity measurement methods: Classification, critique, and implications for the Air Force* (AFHRL-TR-81-9, AD-A105 627). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Tuttle, T.C. (1983). Organizational productivity: A challenge for psychologists. *American Psychologist*, 38, 479-486.

Tuttle, T.C., & Sink, D.S. (1984-85, Winter). Taking the threat out of productivity measurement. *National Productivity Review*, 24-32.

Tuttle, T.C., & Weaver, C.N. (1986). *Methodology for generating efficiency and effectiveness measures (MGEEM): A guide for commanders, managers, and supervisors* (AFHRL-TP-86-26, AD-A167 503). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

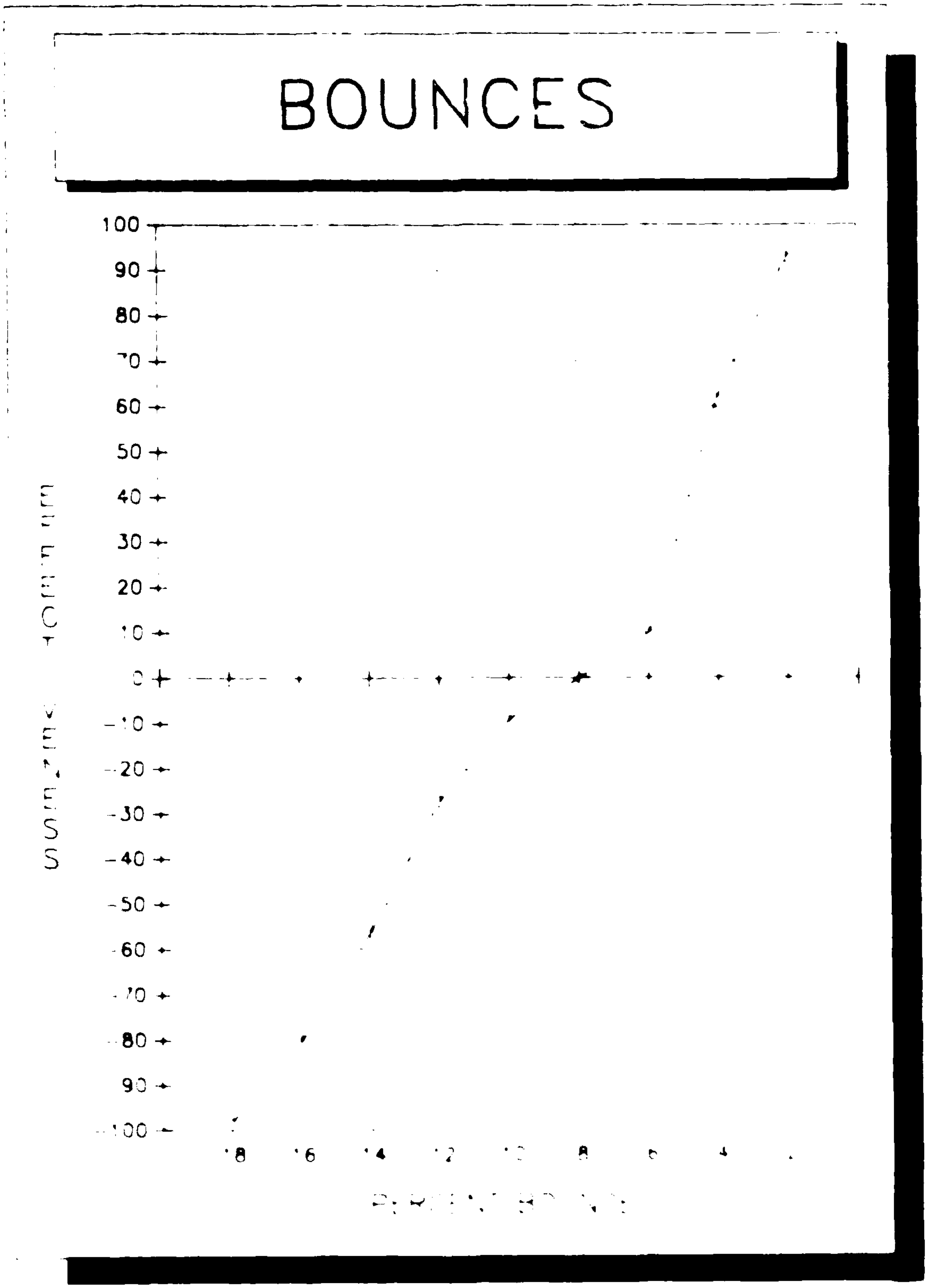
Tuttle, T.C., Wilkinson, R.E., & Matthews, M.D. (1985). *Field test of the methodology for generating efficiency and effectiveness measures* (AFHRL-TP-84-54, AD-A158 183). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

APPENDIX A: CONTINGENCIES

APPENDIX A1: Comm/Nav Contingencies

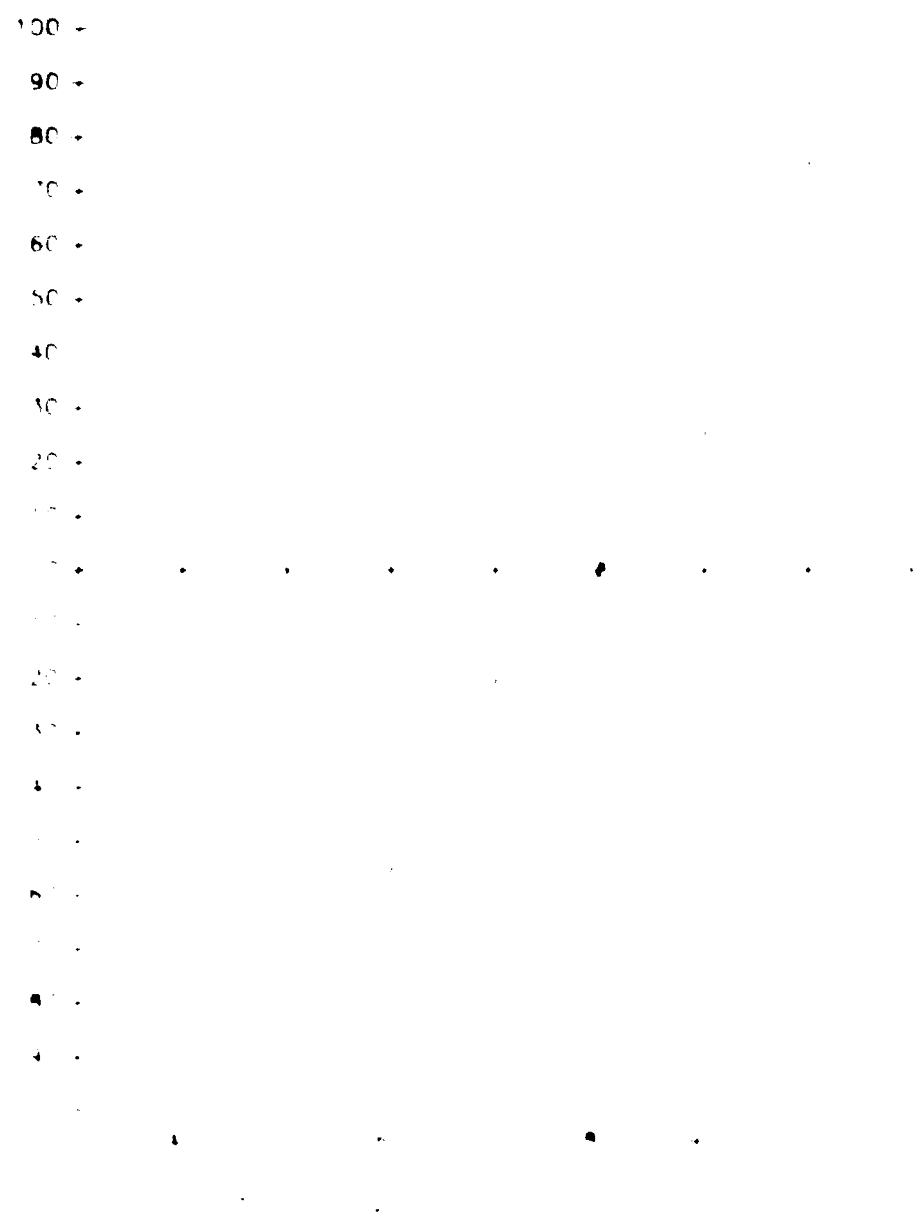
APPENDIX A2: MS&D Contingencies

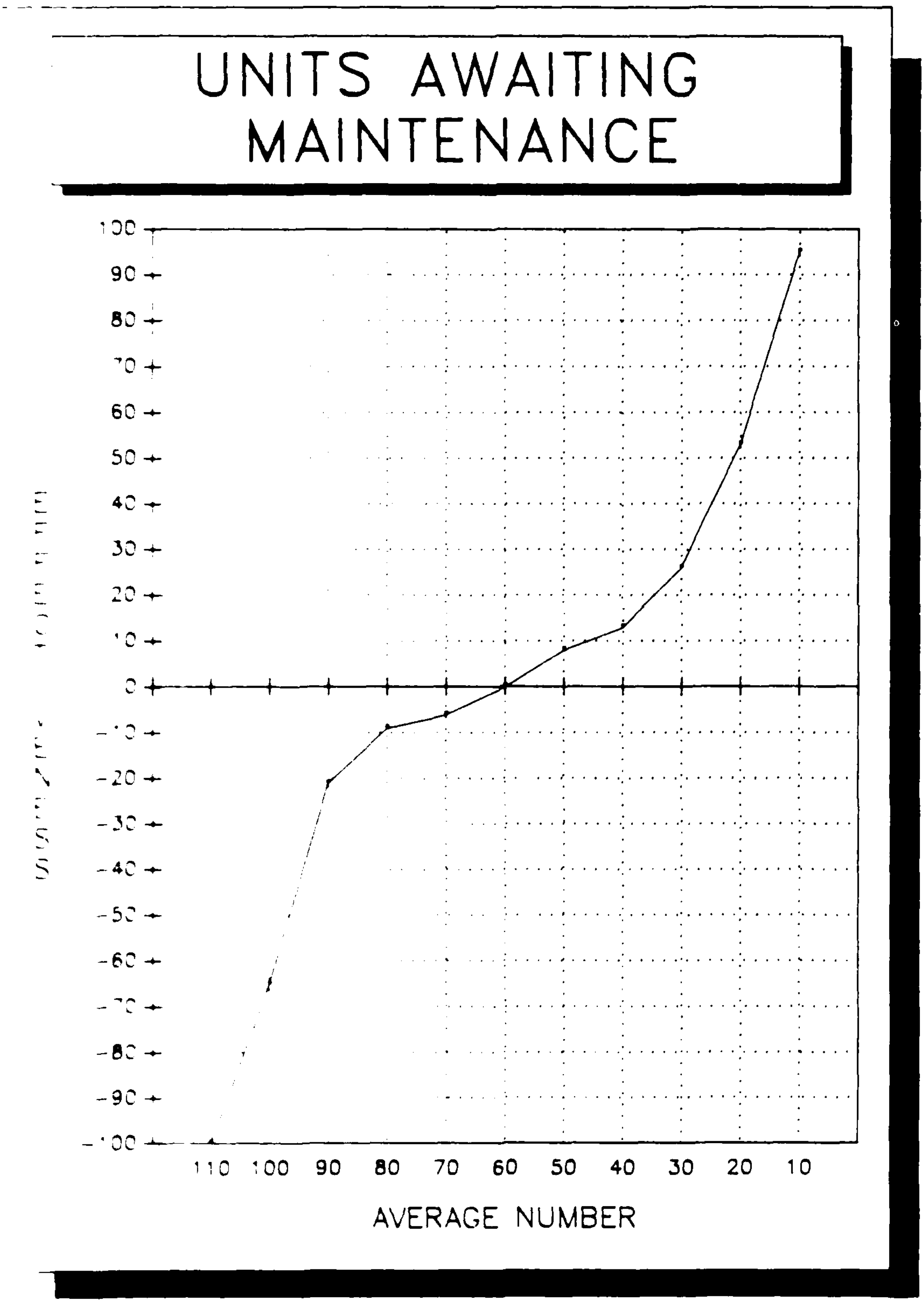
APPENDIX A1: COMM/NAV CONTINGENCIES



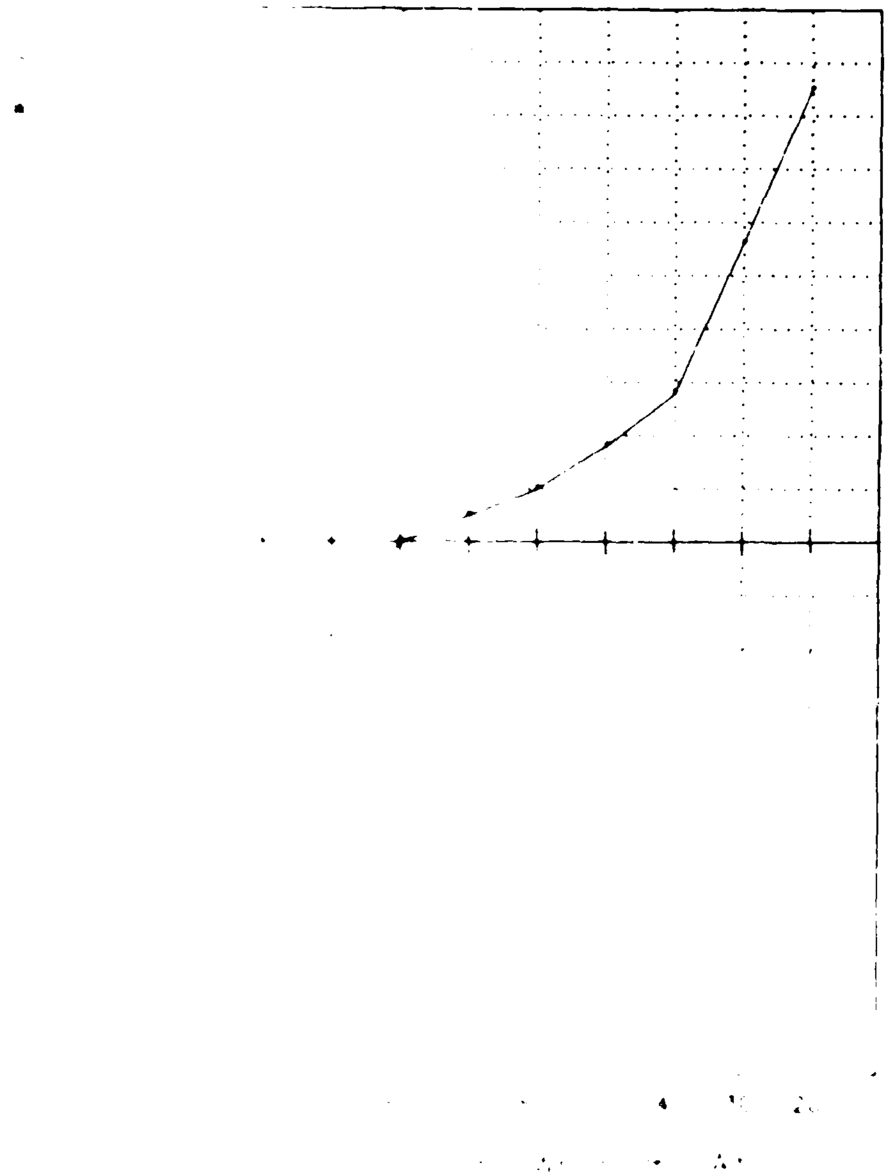
BOUNCES
100
90
80
60
50
40
30
20
0
-20
-30
-40
-50
-60
-80

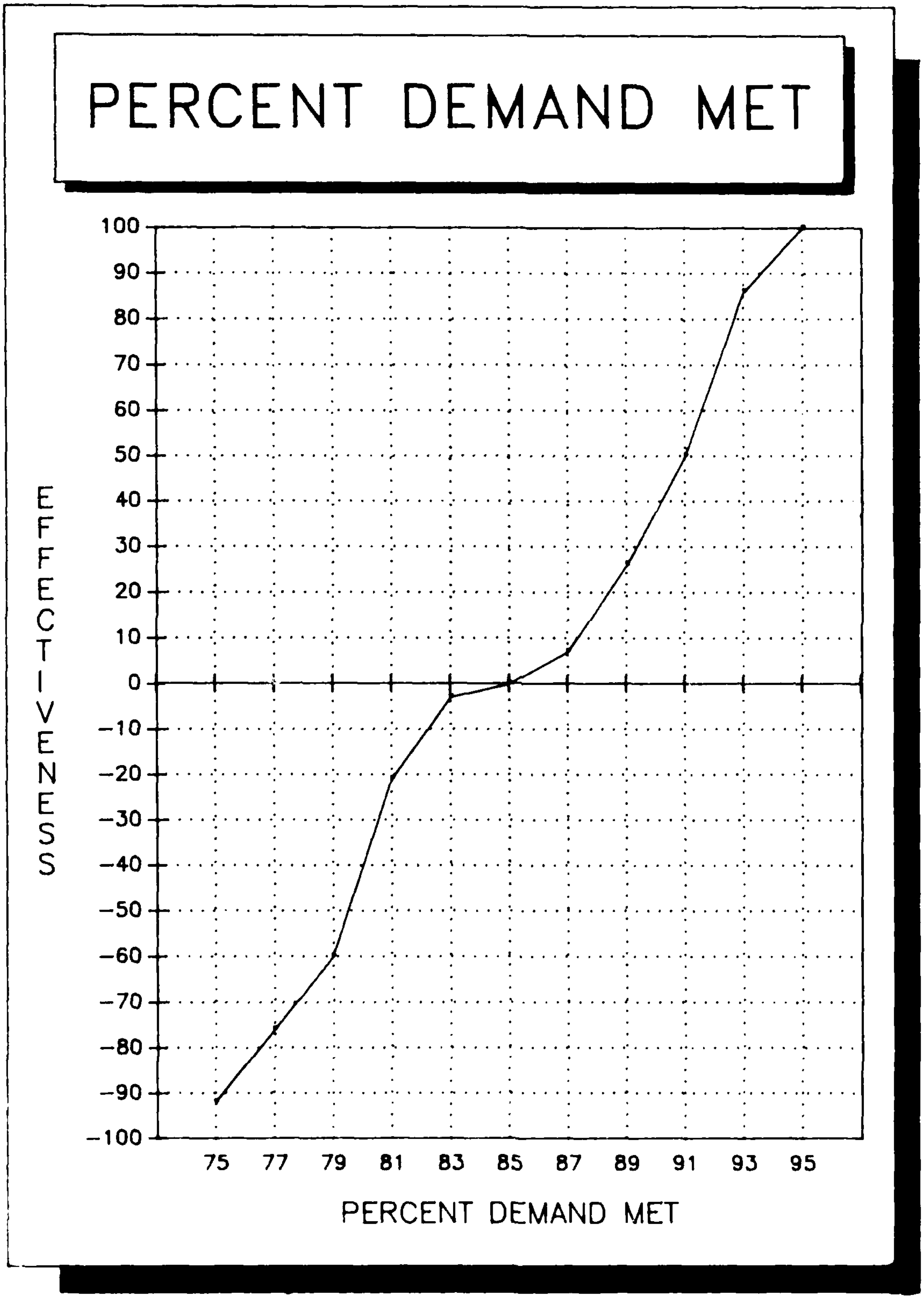
PASSING QA INSPECTIONS

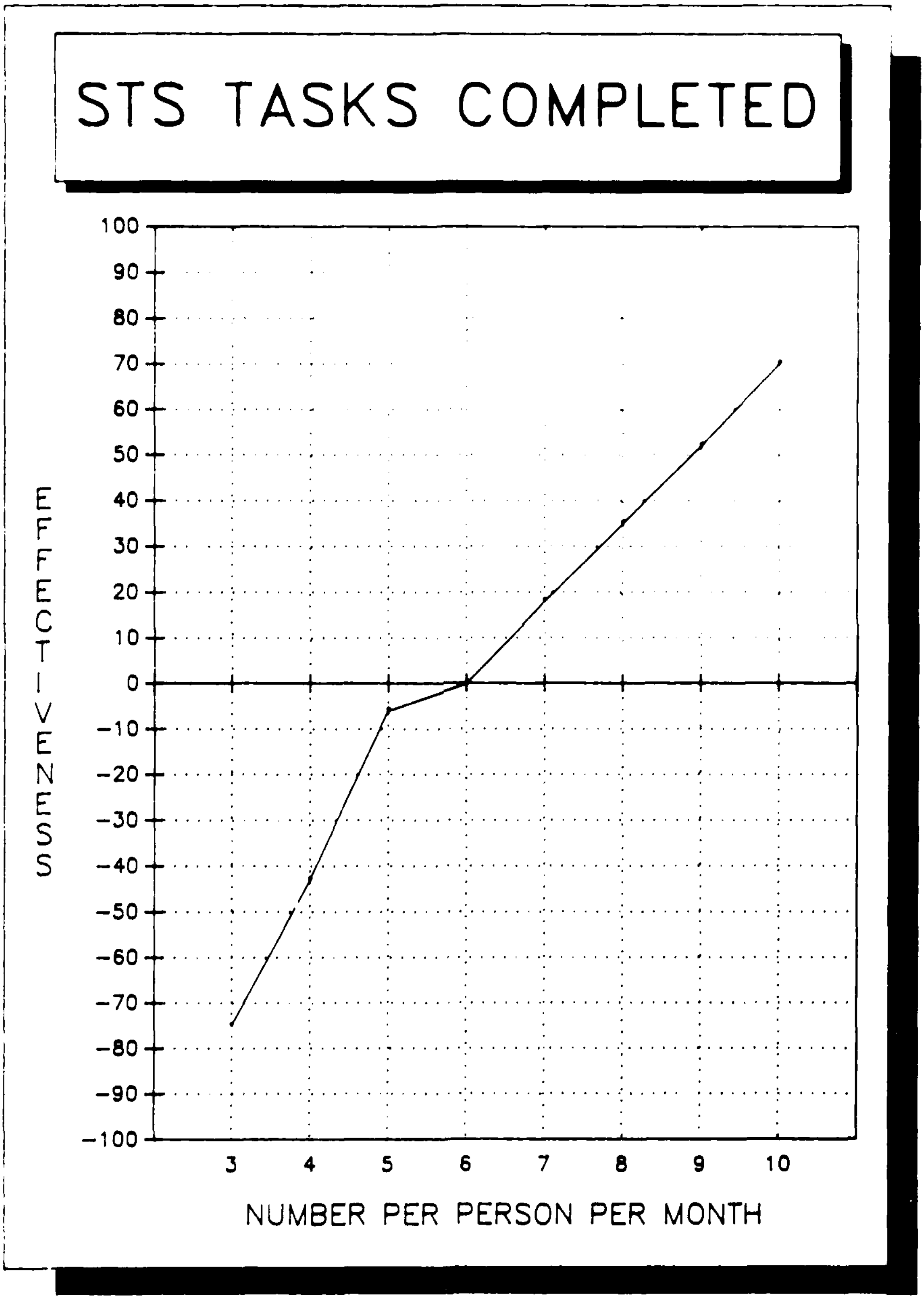



UNITS AWAITING MAINTENANCE
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
110
100
90
80
70
60
50
40
30
20
10
AVERAGE NUMBER

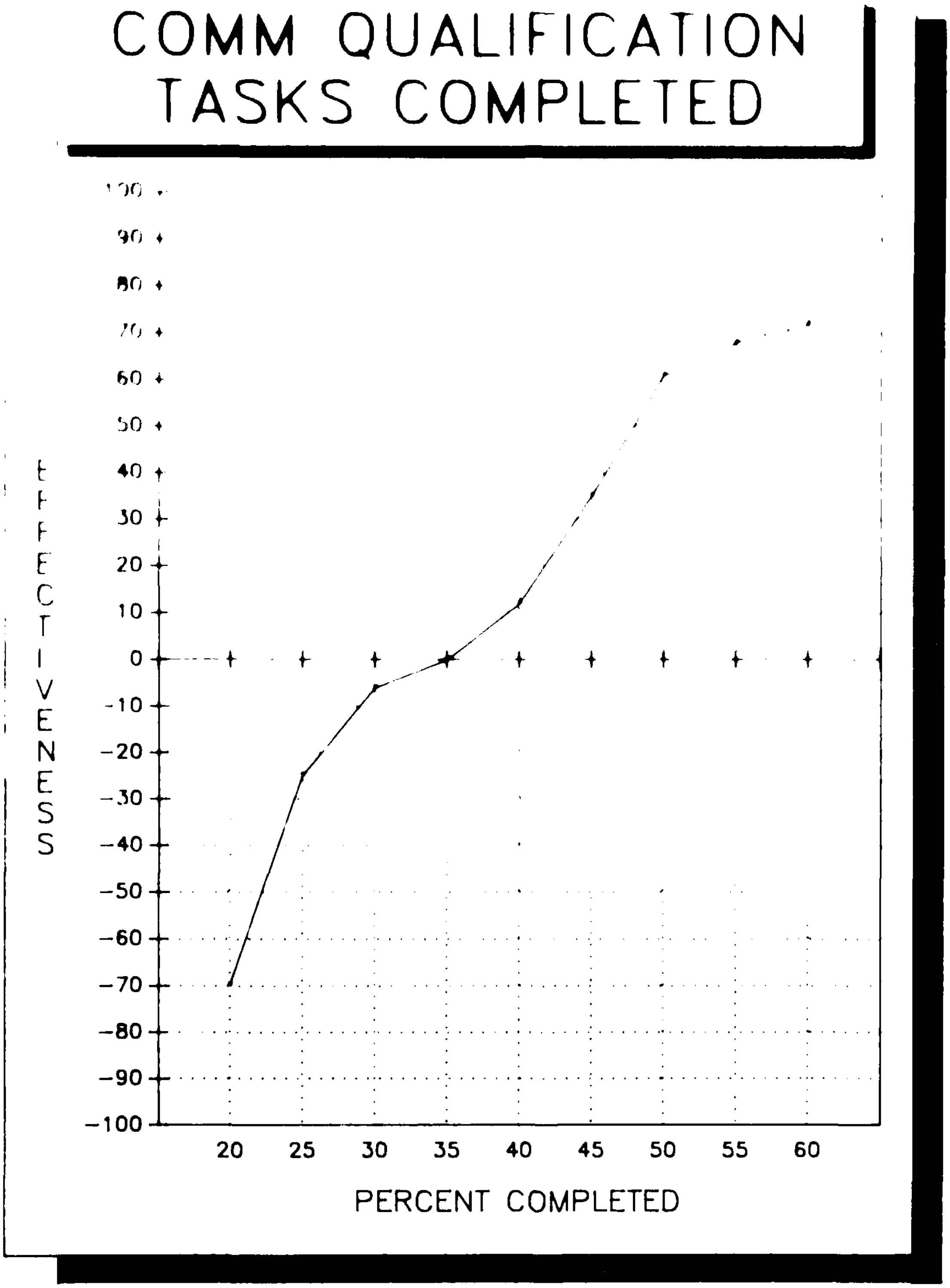
UNITS AWAITING PARTS

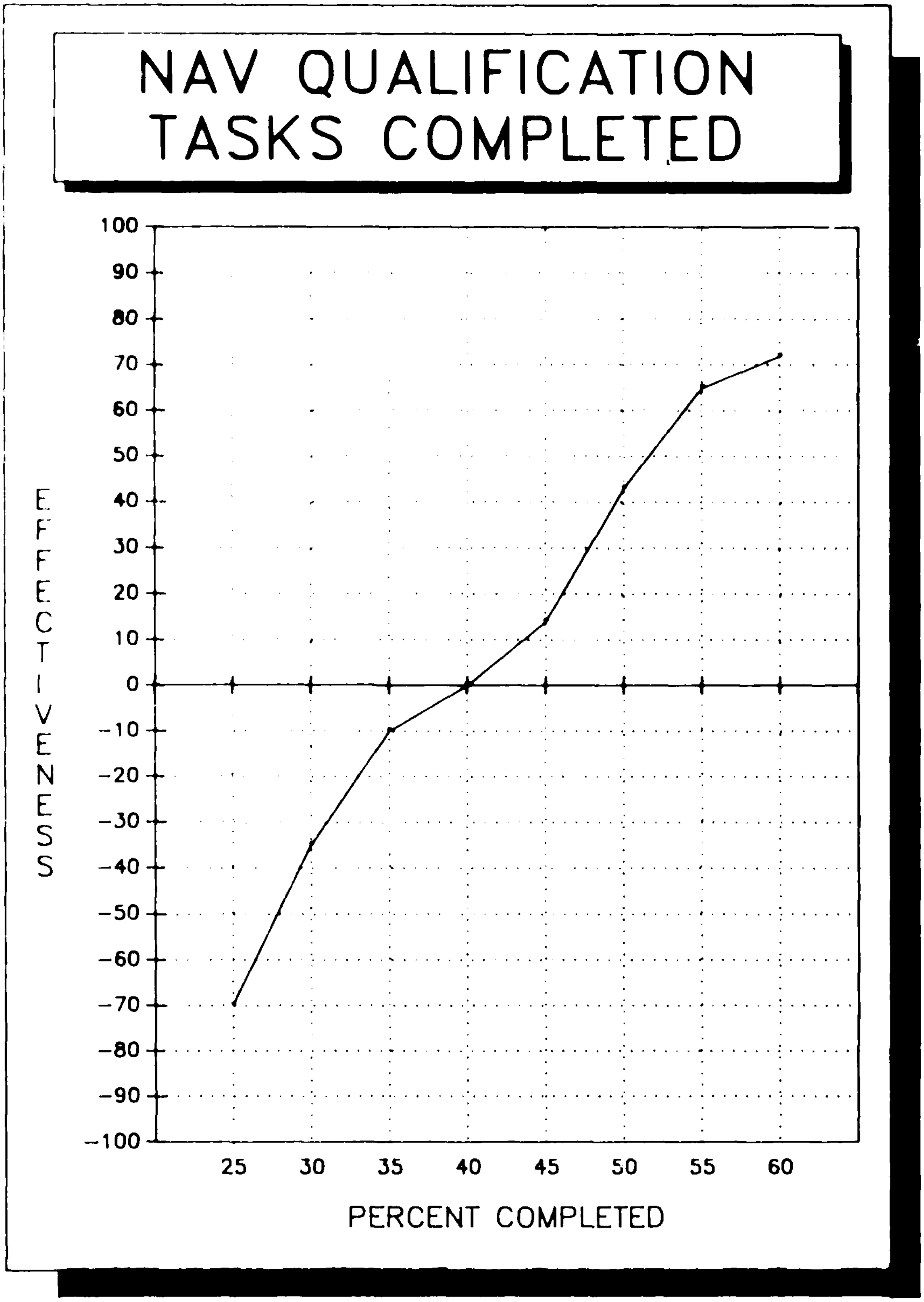


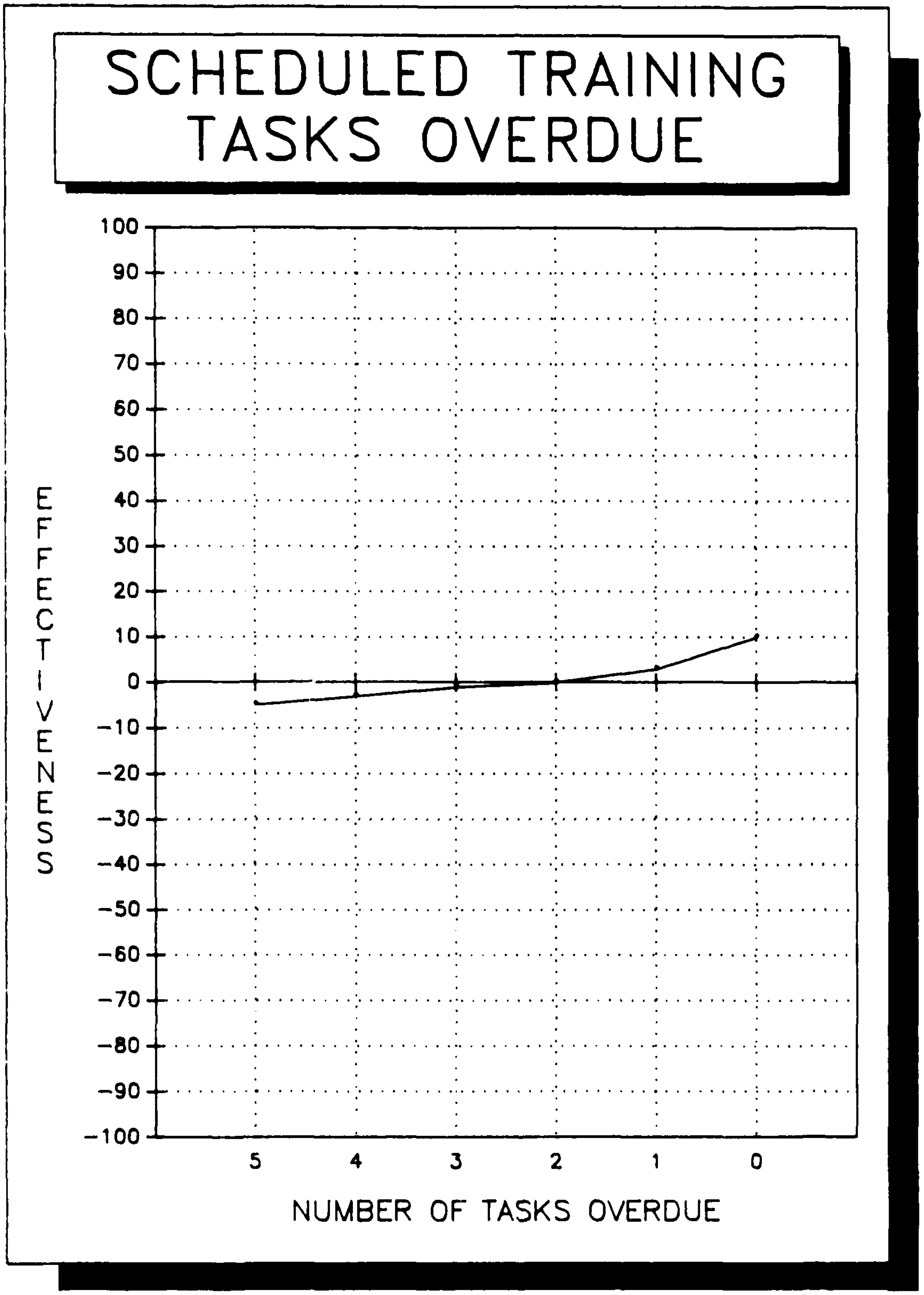

PERCENT DEMAND MET
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
75
77
79
81
83
85
87
89
91
93
95
PERCENT DEMAND MET



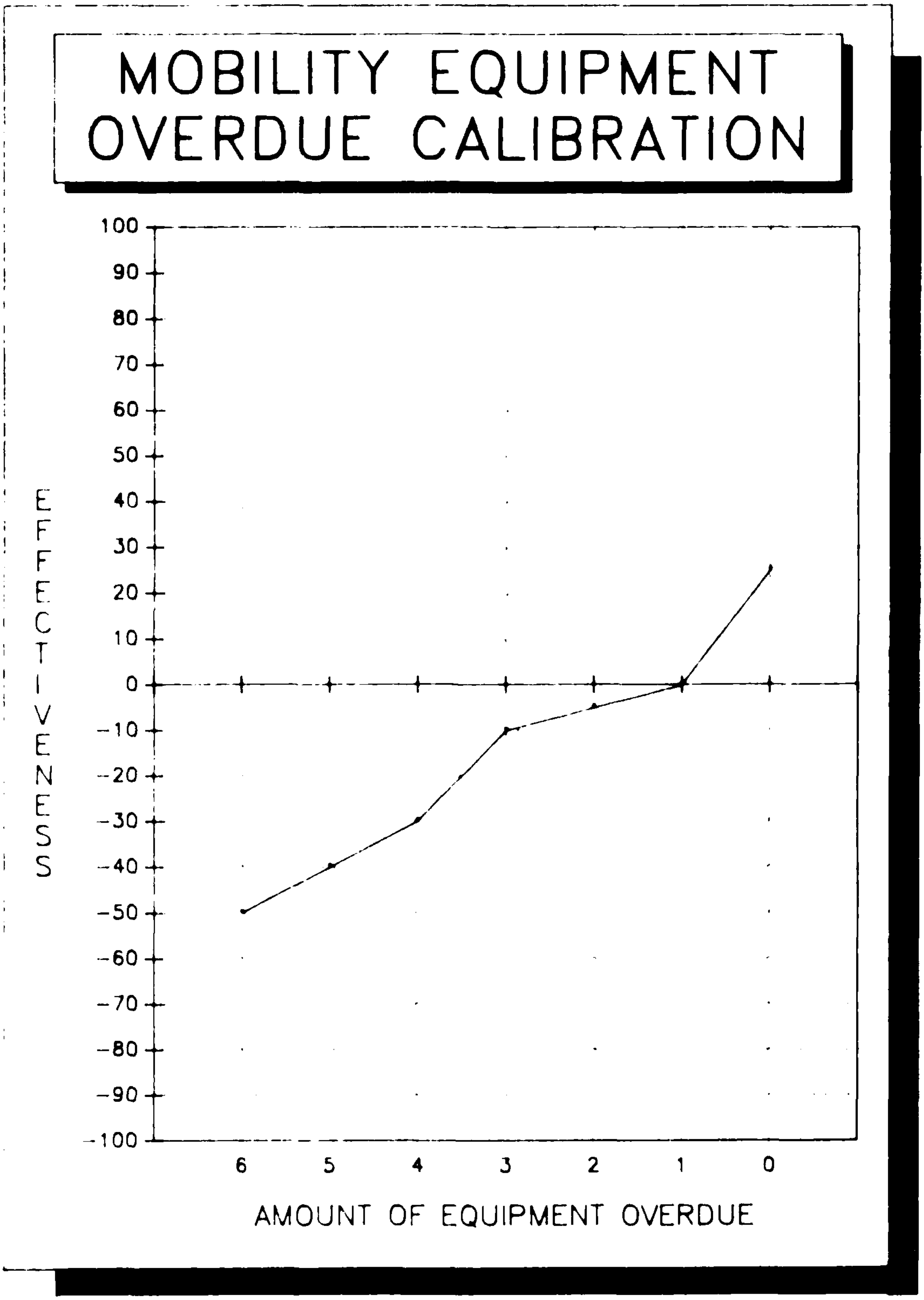
STS TASKS COMPLETED
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−100
3
4
5
6
7
8
9
10
NUMBER PER PERSON PER MONTH

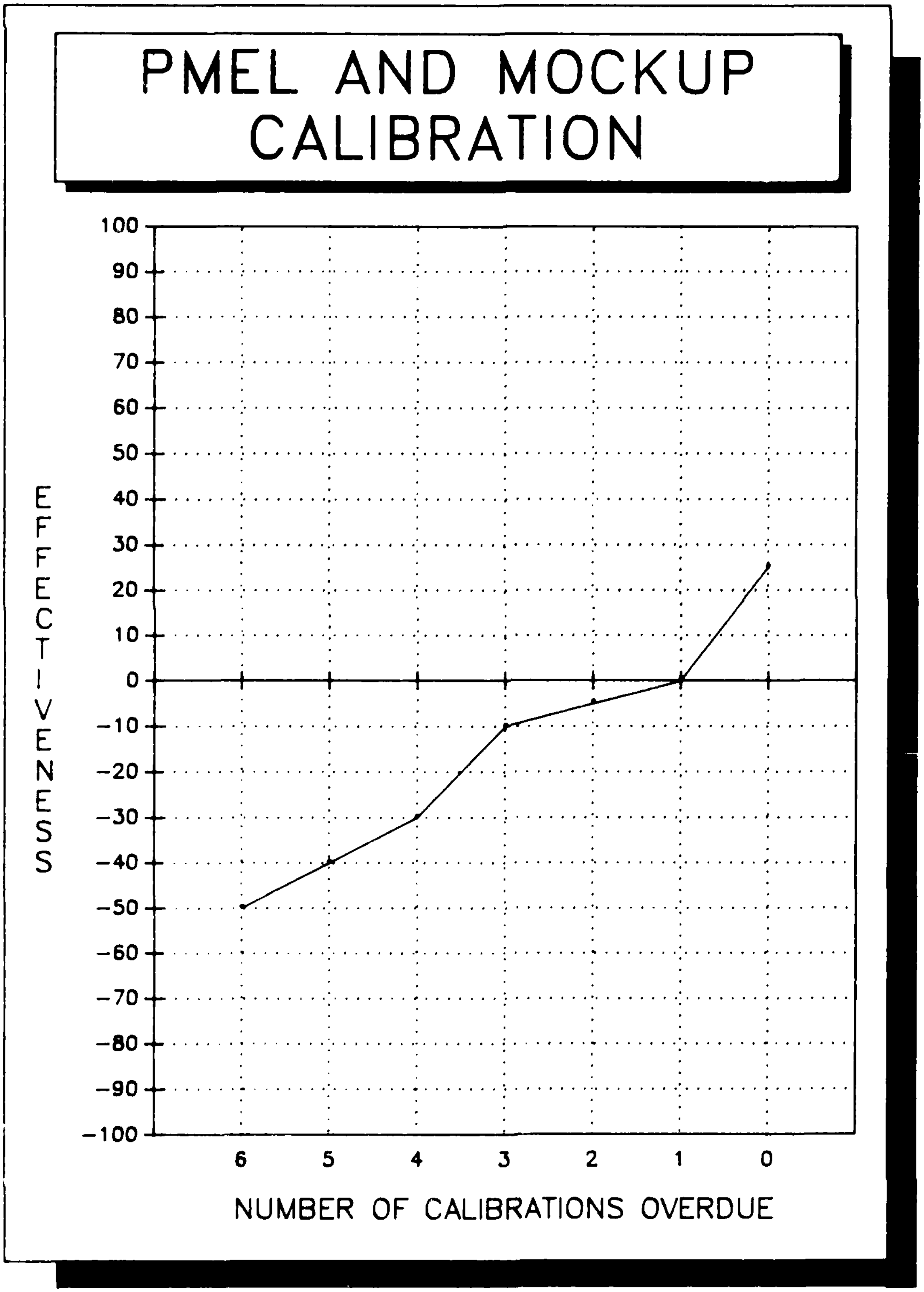

COMM QUALIFICATION
TASKS COMPLETED
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
EFFECTIVENESS
20
25
30
35
40
45
50
55
60
PERCENT COMPLETED

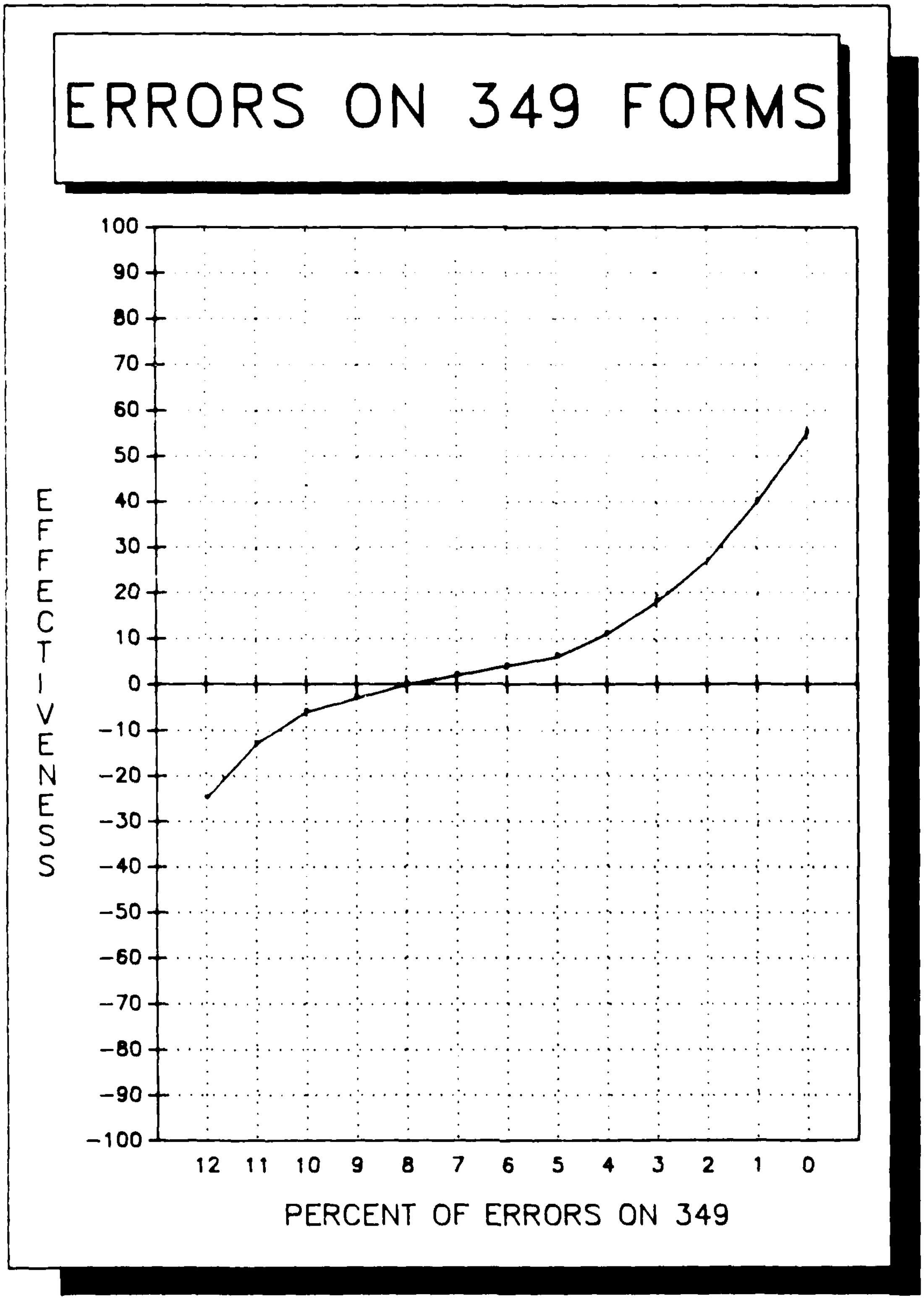

NAV QUALIFICATION
TASKS COMPLETED
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
25
30
35
40
45
50
55
60
PERCENT COMPLETED



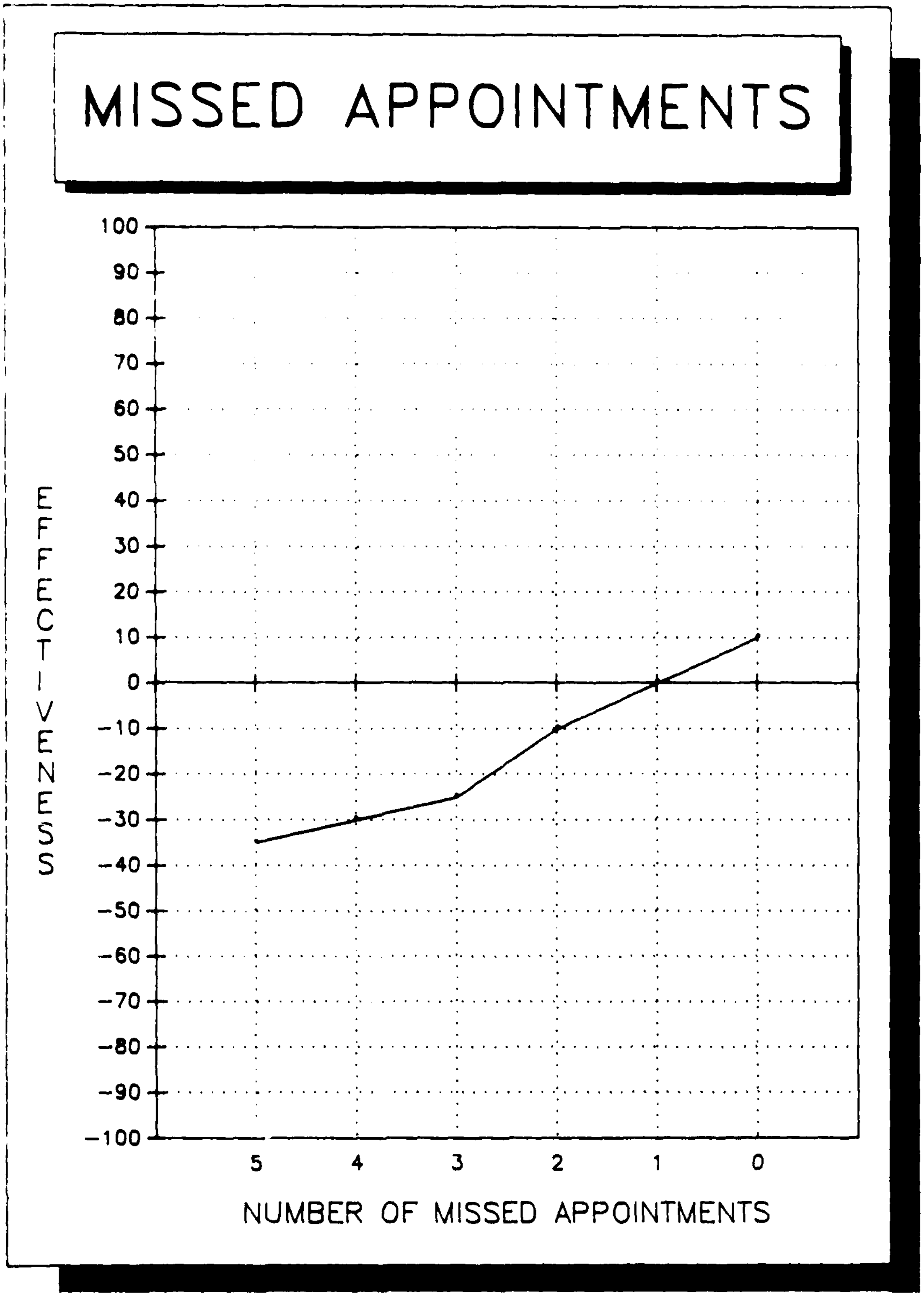
SCHEDULED TRAINING TASKS OVERDUE
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
5
4
3
2
1
0
NUMBER OF TASKS OVERDUE


MOBILITY EQUIPMENT
OVERDUE CALIBRATION
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
6
5
4
3
2
1
0
AMOUNT OF EQUIPMENT OVERDUE

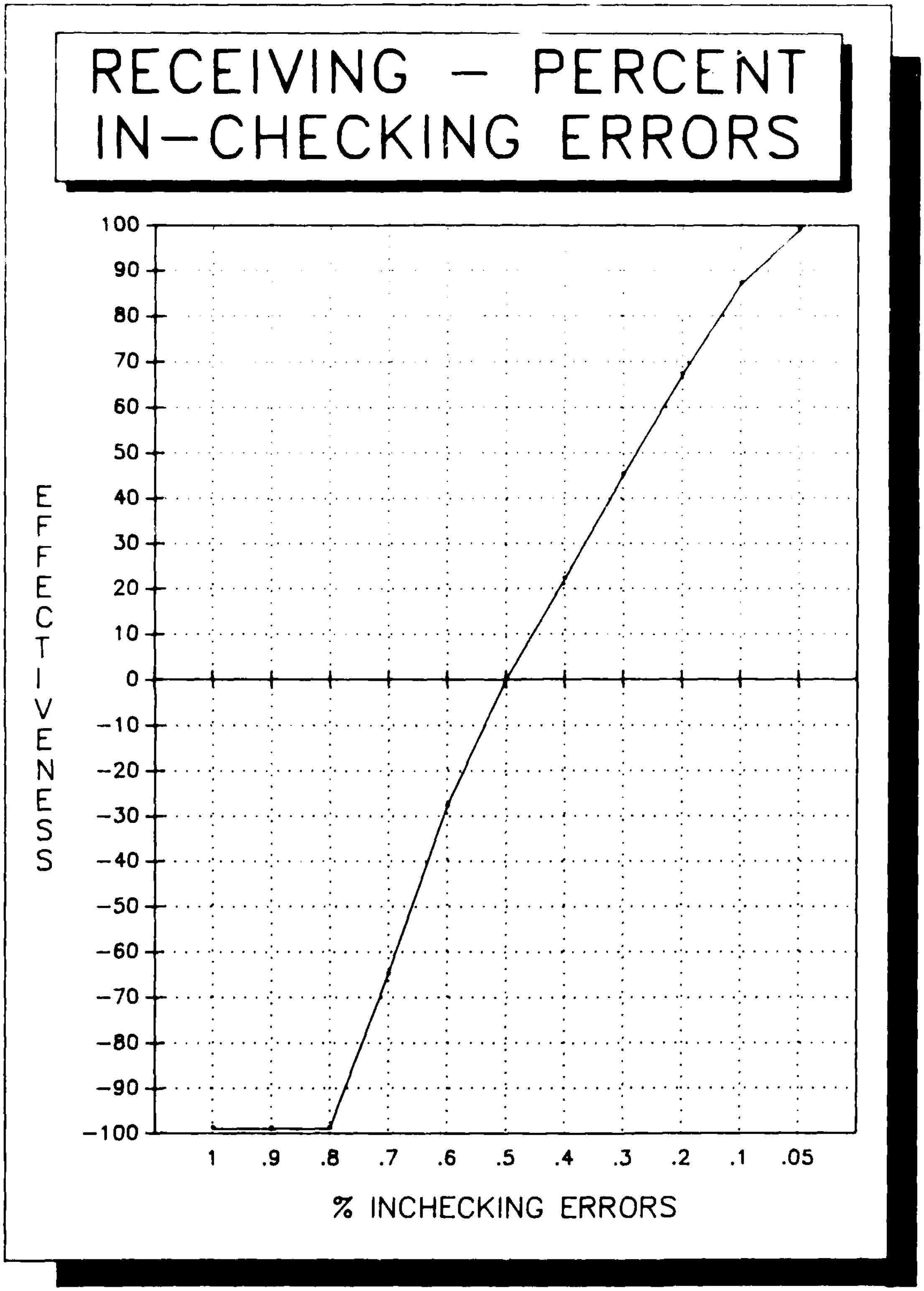

PMEL AND MOCKUP
CALIBRATION
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
6
5
4
3
2
1
0
NUMBER OF CALIBRATIONS OVERDUE

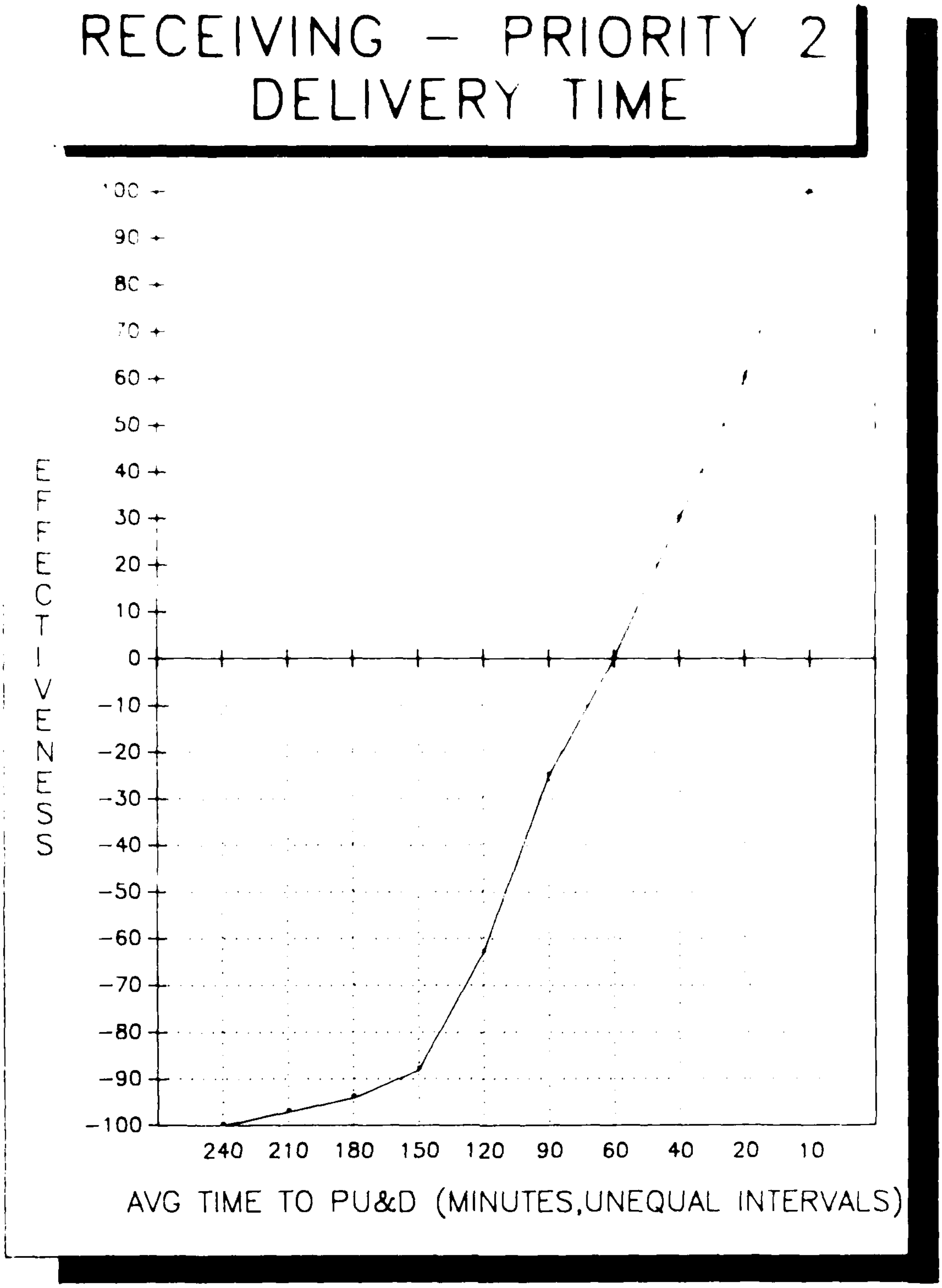


ERRORS ON 349 FORMS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
12 11 10 9 8 7 6 5 4 3 2 1 0
PERCENT OF ERRORS ON 349

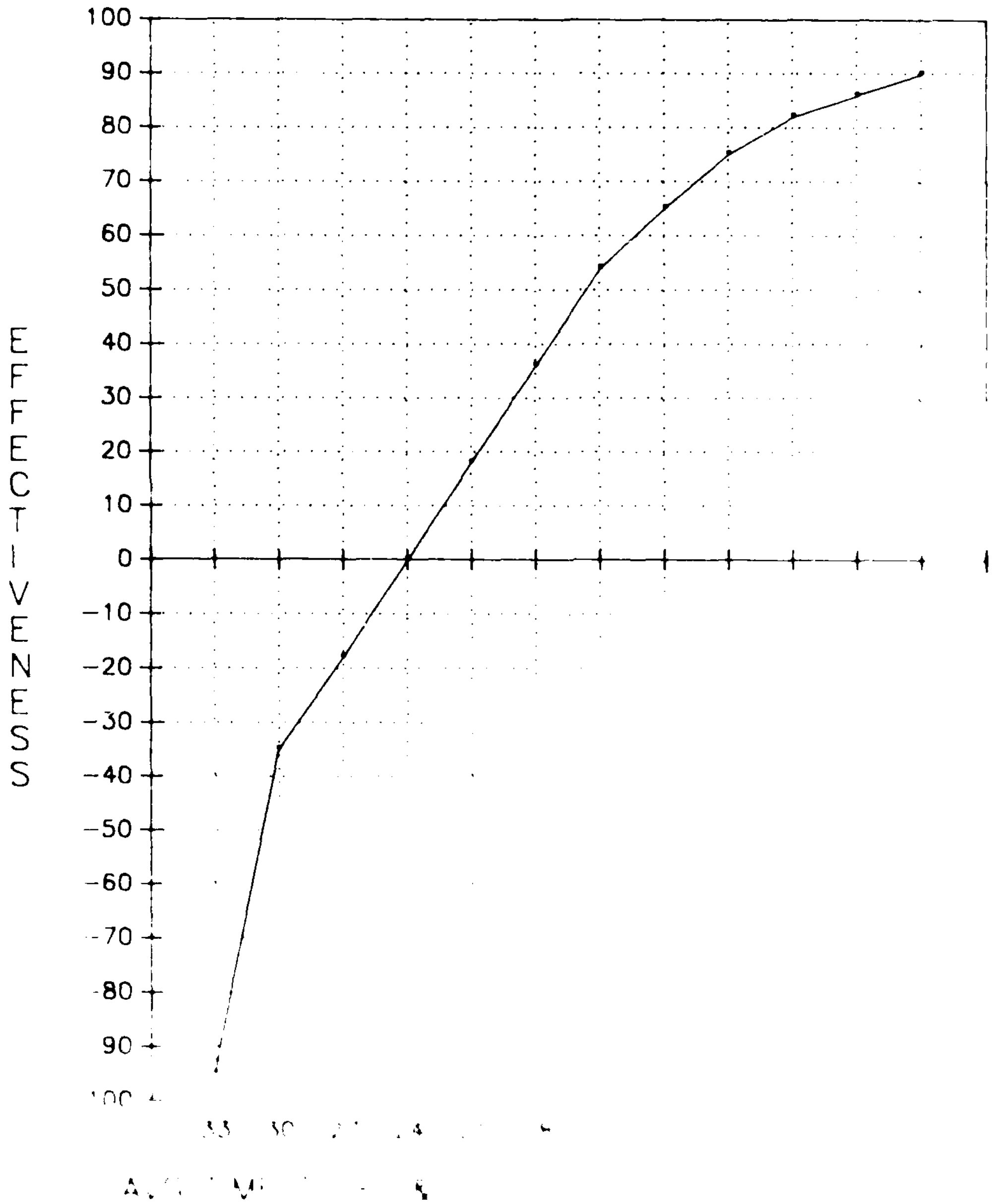

MISSED APPOINTMENTS
100
90
80
70
60
50
40
30
20
10
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−100
EFFECTIVENESS
5
4
3
2
1
0
NUMBER OF MISSED APPOINTMENTS

APPENDIX A2: MS&D CONTINGENCIES


RECEIVING — PERCENT IN-CHECKING ERRORS
100
90
80
70
60
50
40
30
20
10
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−100
EFFECTIVENESS
1
.9
.8
.7
.6
.5
.4
.3
.2
.1
.05
% INCHECKING ERRORS


RECEIVING — PRIORITY 2
DELIVERY TIME
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
240 210 180 150 120 90 60 40 20 10
AVG TIME TO PU&D (MINUTES,UNEQUAL INTERVALS)

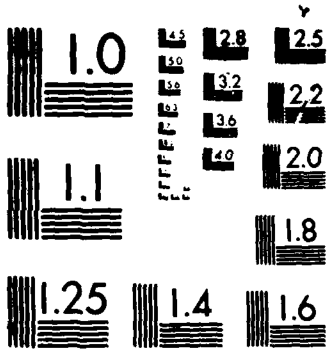
RECEIVING – PRIORITY 4
DELIVERY TIME



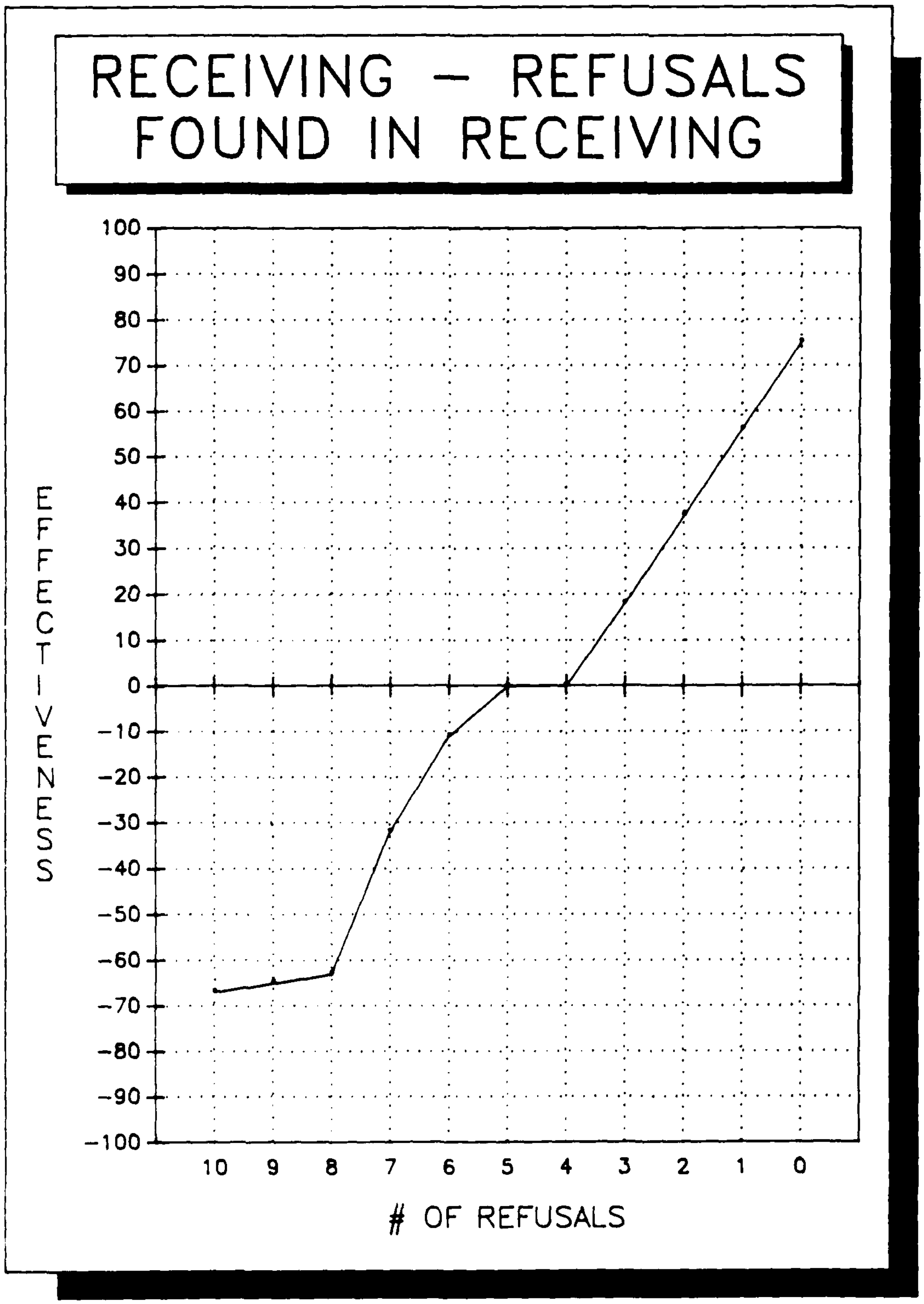
AD-A183 565 ORGANIZATIONAL PRODUCTIVITY MEASUREMENT: THE DEVELOPMENT AND EVALUATION O.. (U) HOUSTON UNIV TX DEPT OF PSYCHOLOGY R D PRITCHARD ET AL. JUL 87 2/2

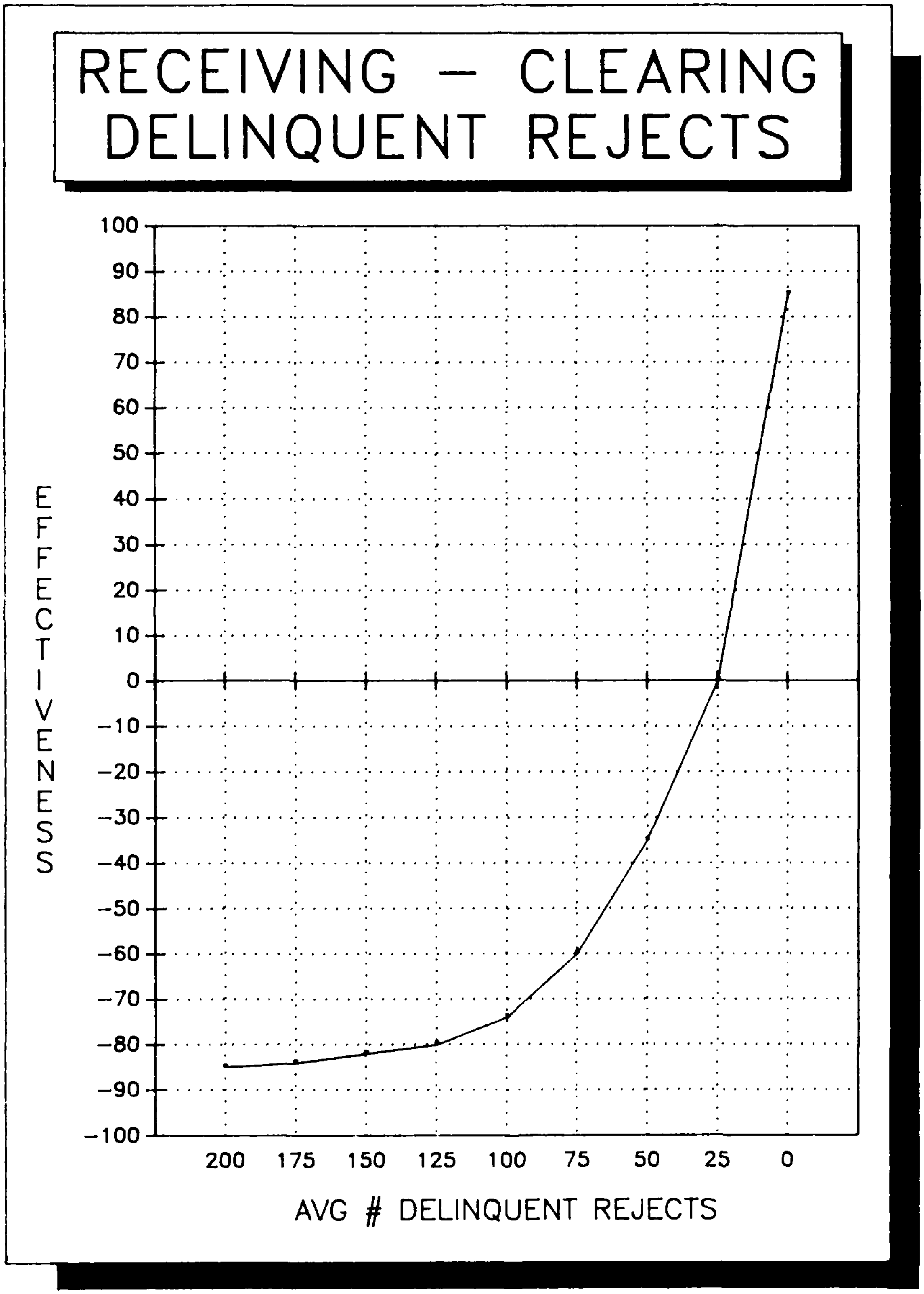
UNCLASSIFIED AFHRL-TR-86-64 F41689-83-C-0039 F/G 5/1 NL

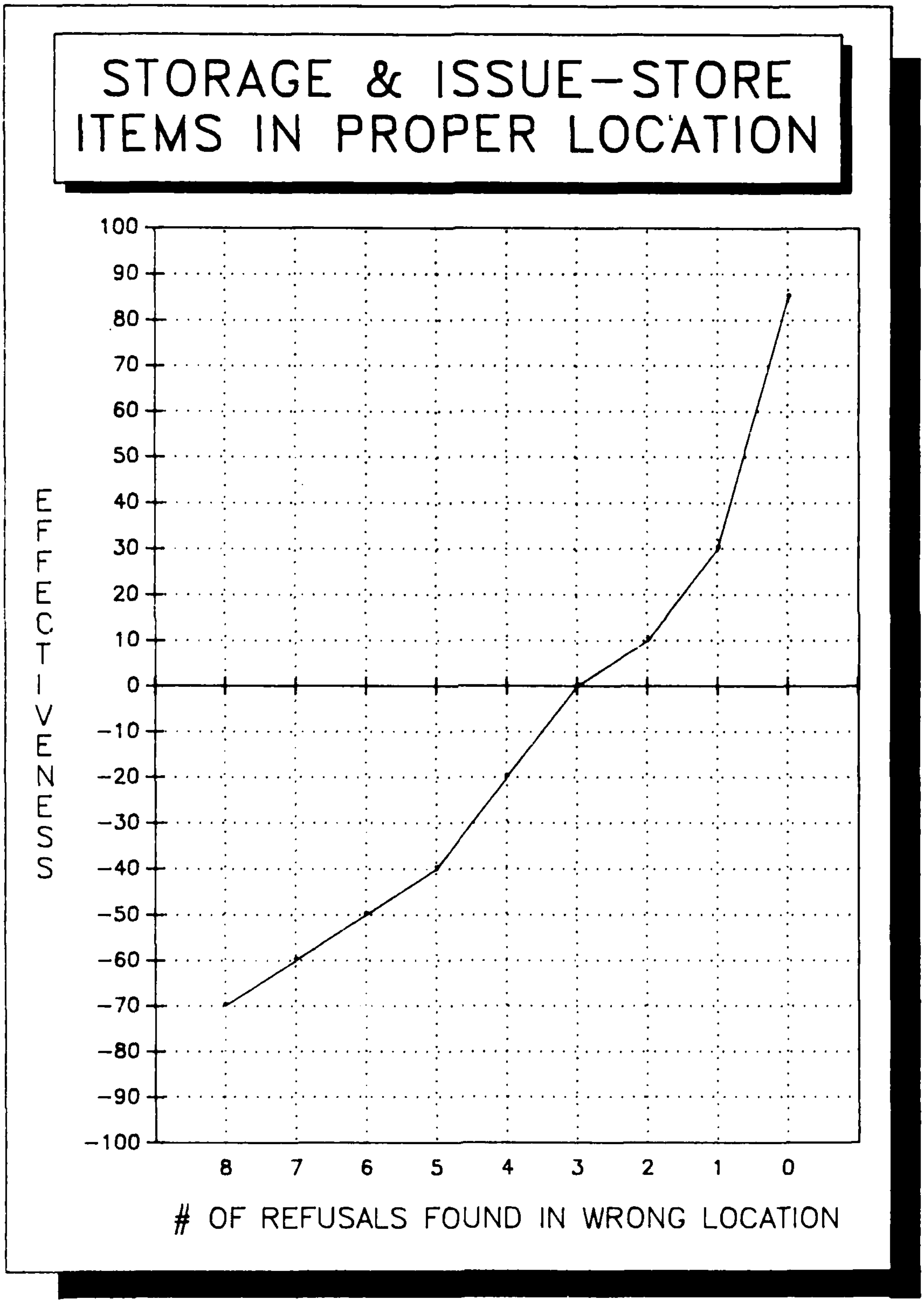
END
9-87
DTIC

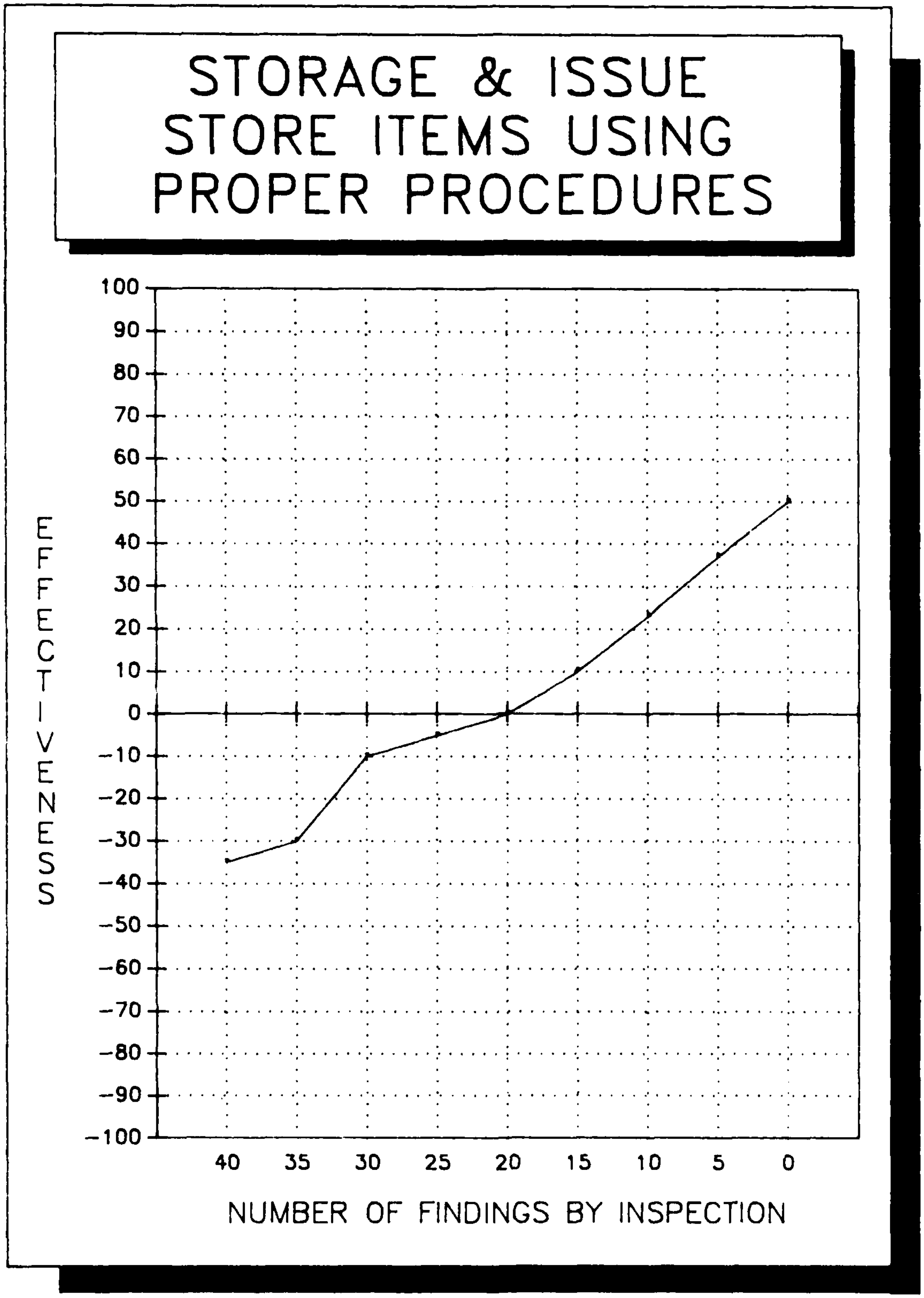


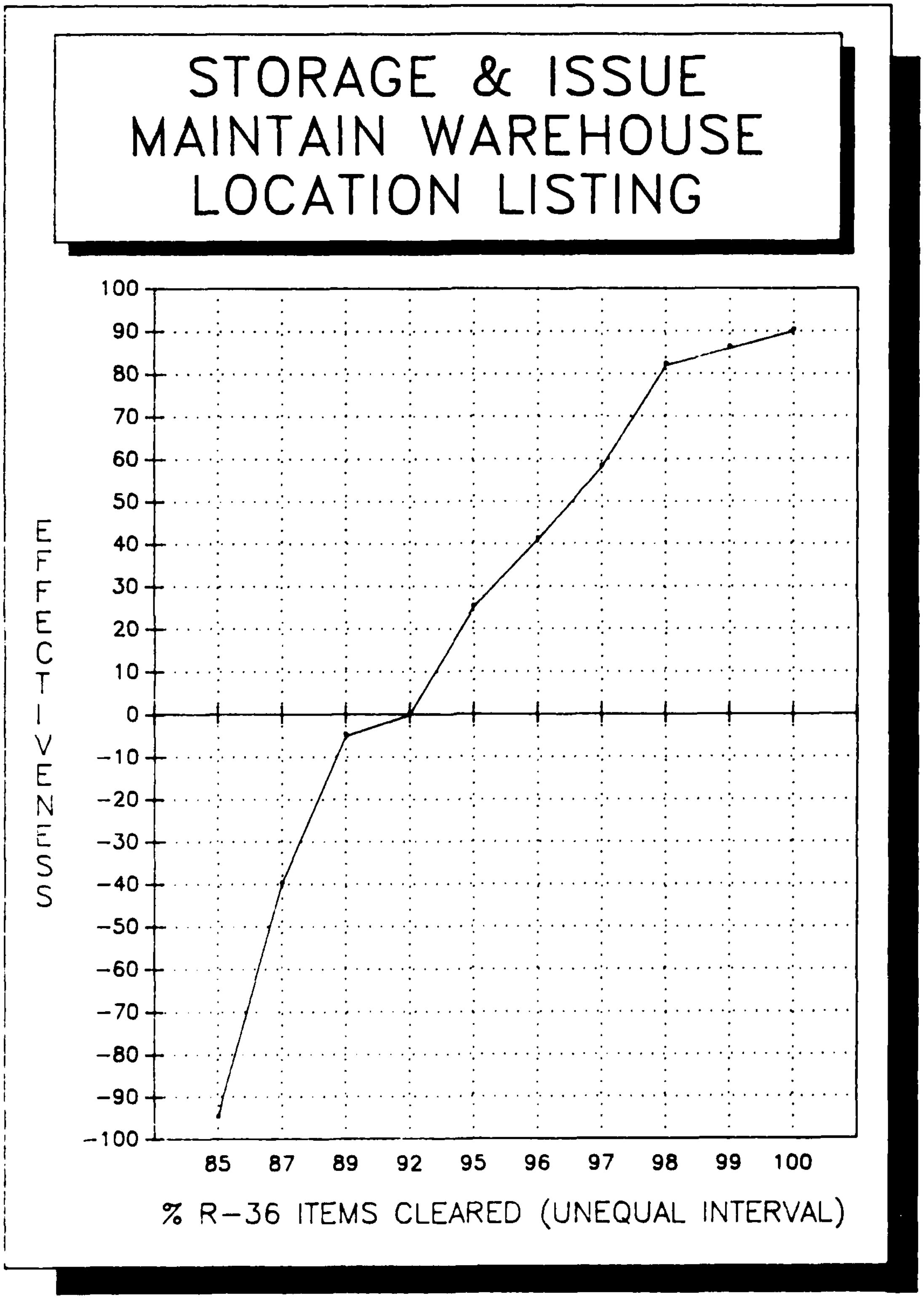
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


RECEIVING - REFUSALS
FOUND IN RECEIVING
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
10
9
8
7
6
5
4
3
2
1
0
OF REFUSALS

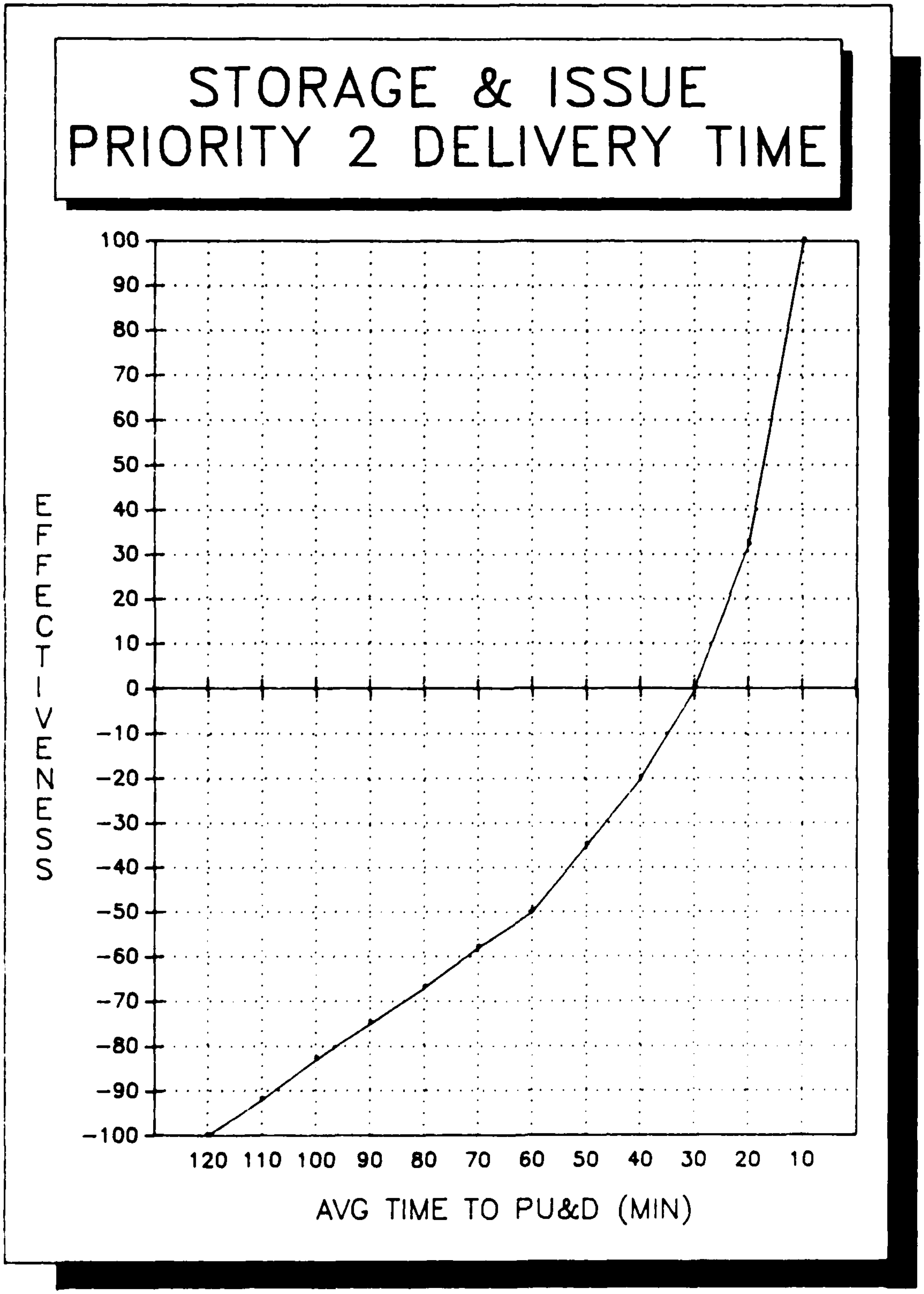

RECEIVING – CLEARING
DELINQUENT REJECTS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−100
200
175
150
125
100
75
50
25
0
AVG # DELINQUENT REJECTS

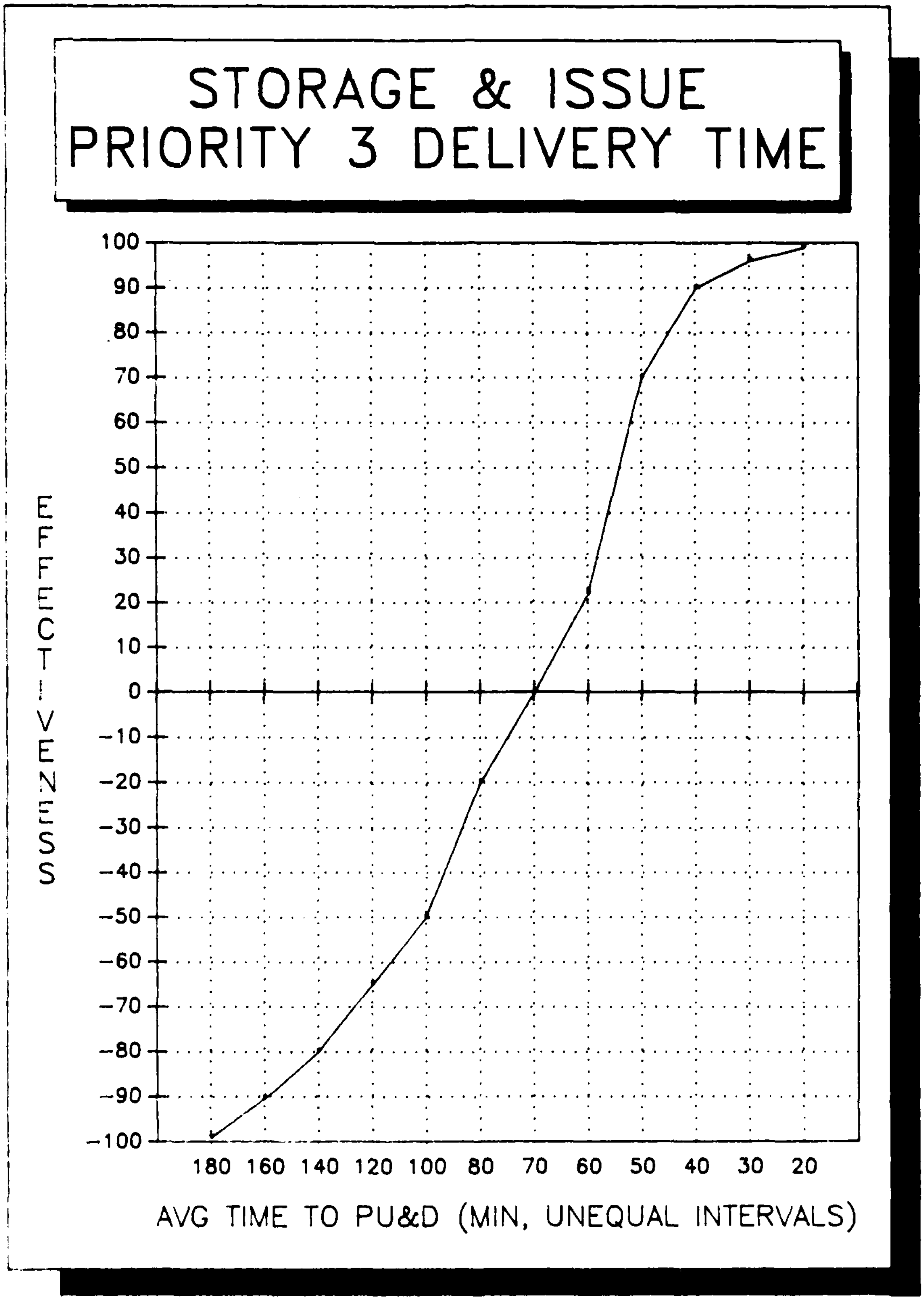

STORAGE & ISSUE-STORE
ITEMS IN PROPER LOCATION
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
8
7
6
5
4
3
2
1
0
OF REFUSALS FOUND IN WRONG LOCATION


STORAGE & ISSUE
STORE ITEMS USING
PROPER PROCEDURES
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−100
40
35
30
25
20
15
10
5
0
NUMBER OF FINDINGS BY INSPECTION

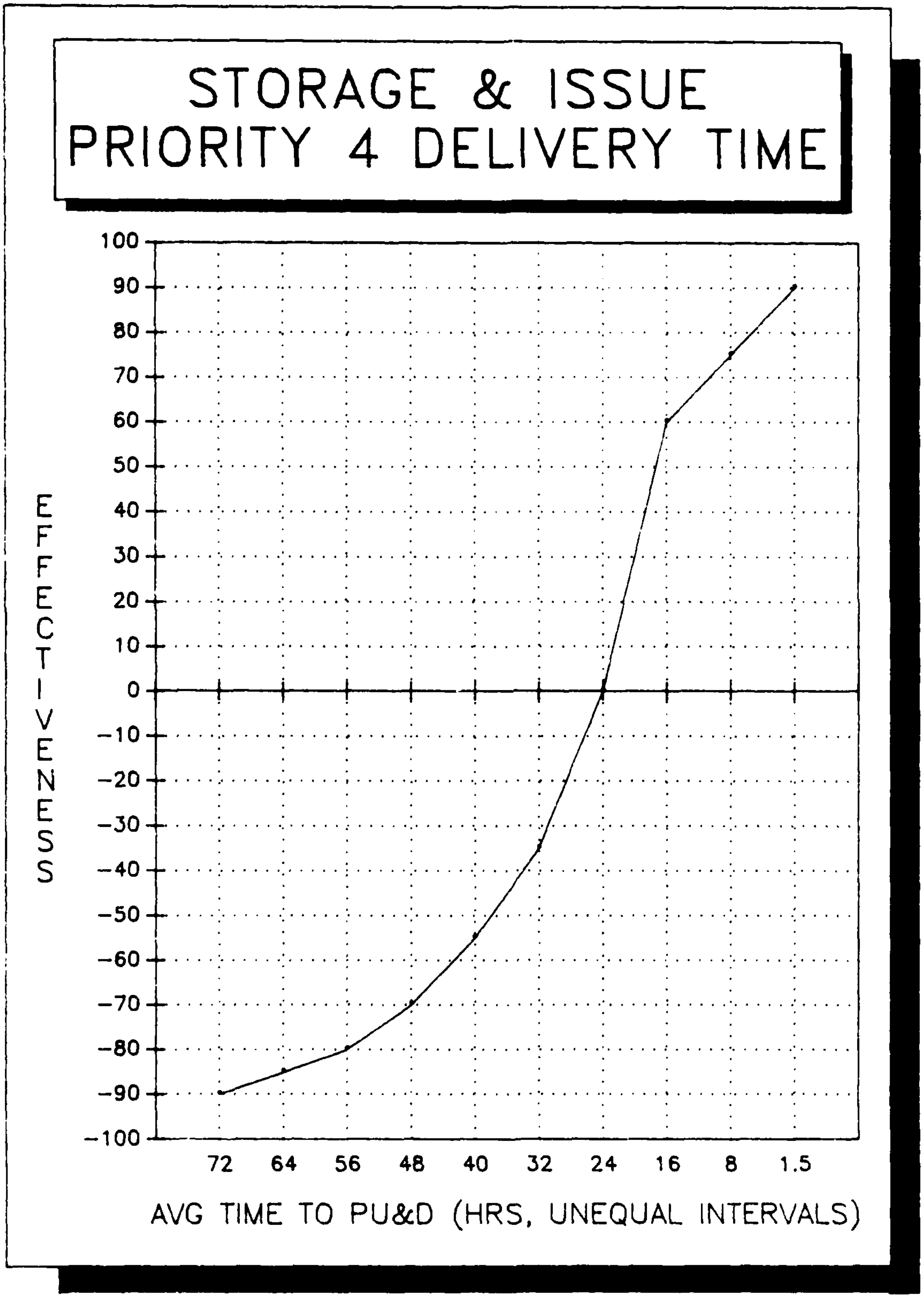


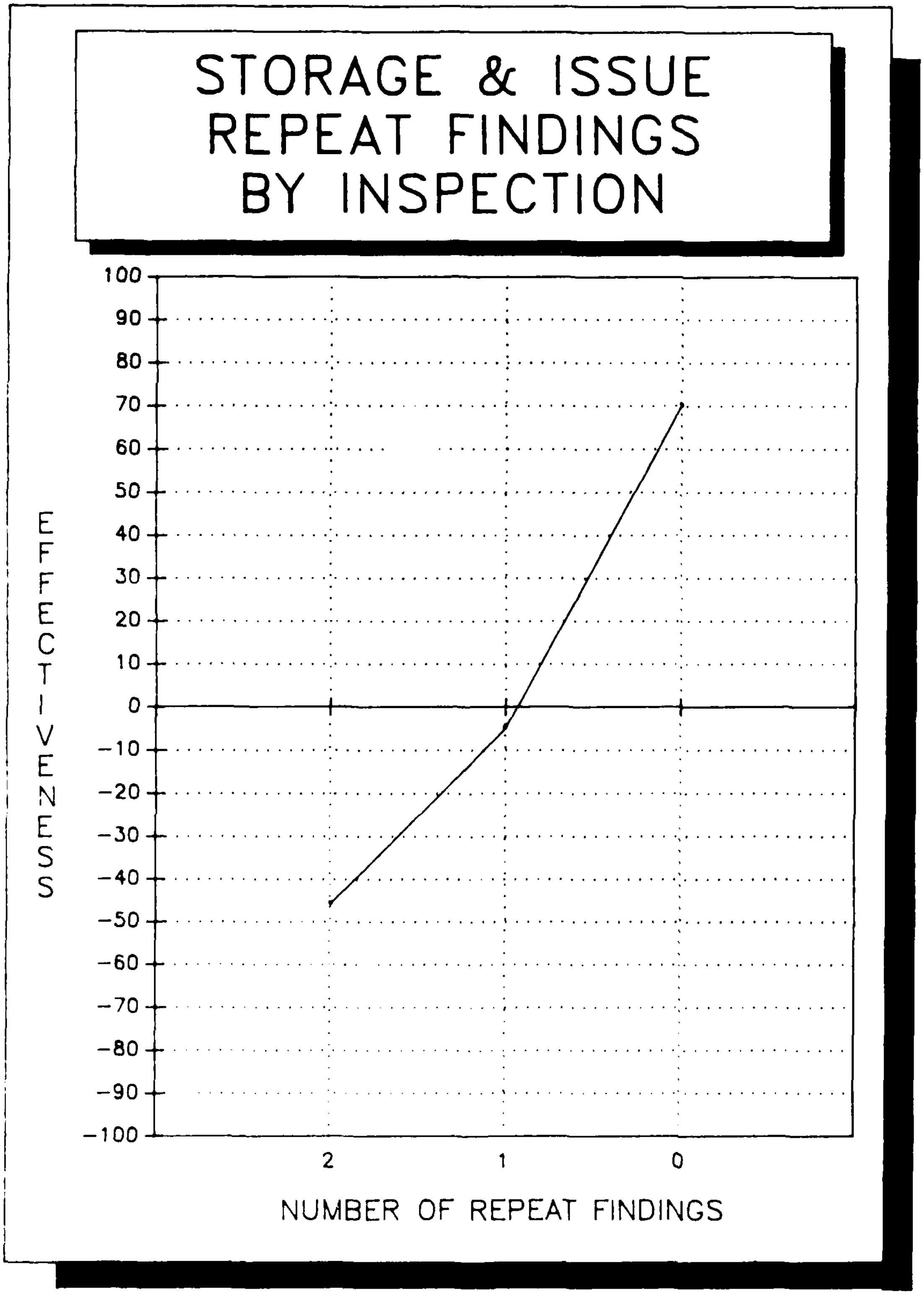
STORAGE & ISSUE
MAINTAIN WAREHOUSE
LOCATION LISTING
100
90
80
70
60
50
40
30
20
10
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−100
EFFECTIVENESS
85
87
89
92
95
96
97
98
99
100
% R−36 ITEMS CLEARED (UNEQUAL INTERVAL)

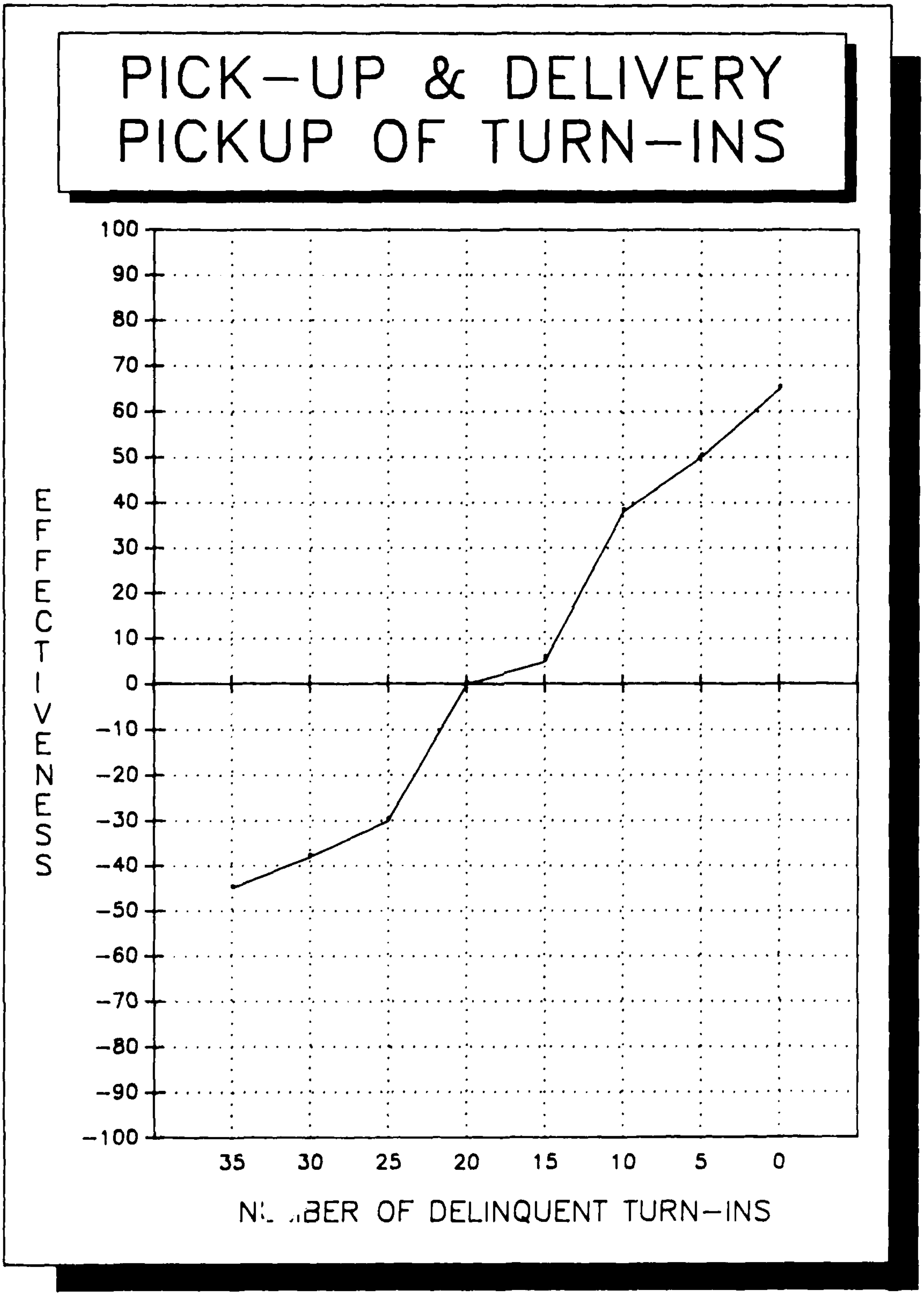

STORAGE & ISSUE
PRIORITY 2 DELIVERY TIME
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−100
120 110 100 90 80 70 60 50 40 30 20 10
AVG TIME TO PU&D (MIN)



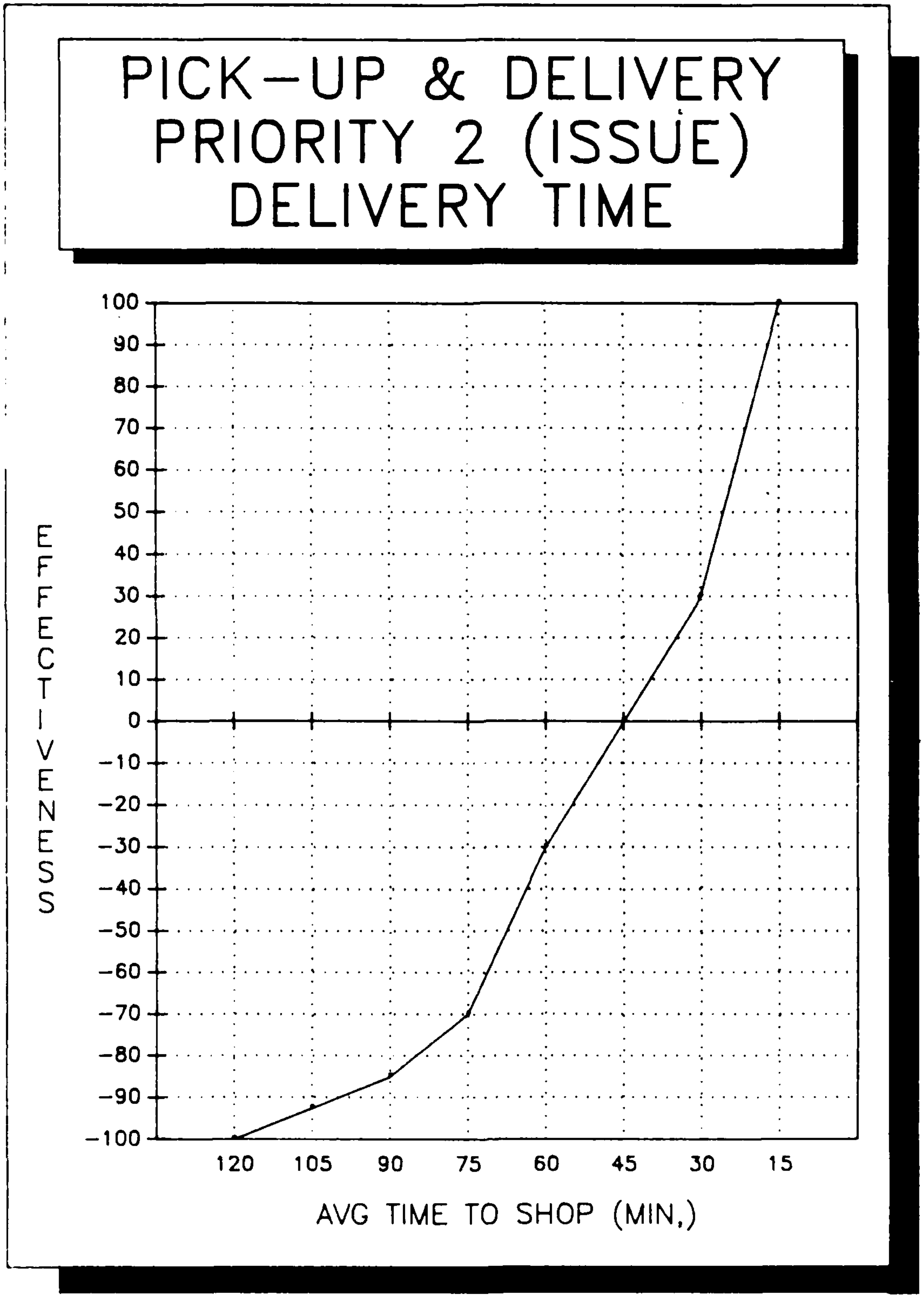
STORAGE & ISSUE
PRIORITY 3 DELIVERY TIME
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
180 160 140 120 100 80 70 60 50 40 30 20
AVG TIME TO PU&D (MIN, UNEQUAL INTERVALS)


STORAGE & ISSUE
PRIORITY 4 DELIVERY TIME
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
EFFECTIVENESS
72
64
56
48
40
32
24
16
8
1.5
AVG TIME TO PU&D (HRS, UNEQUAL INTERVALS)

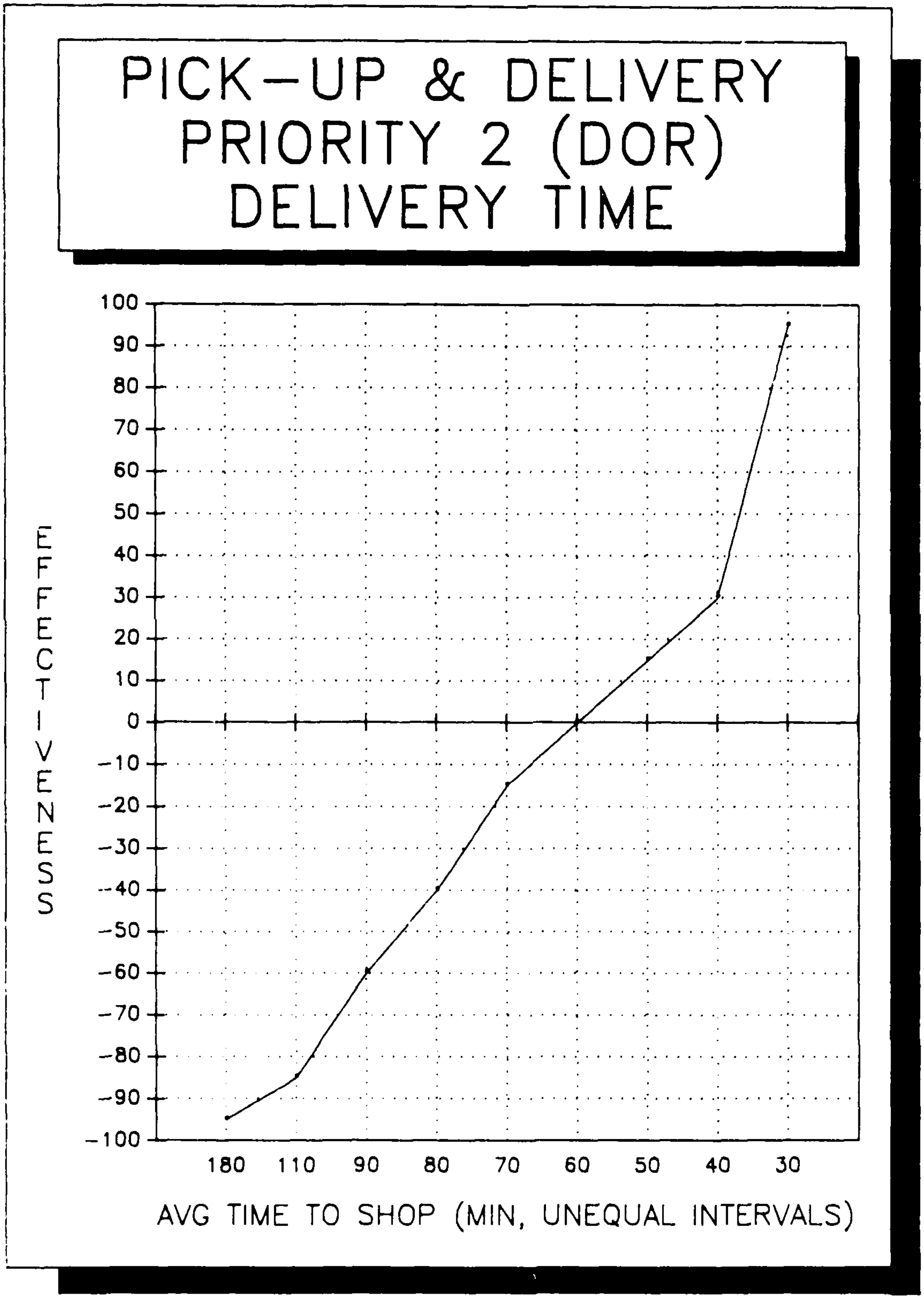

STORAGE & ISSUE
REPEAT FINDINGS
BY INSPECTION
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
EFFECTIVENESS
2
1
0
NUMBER OF REPEAT FINDINGS

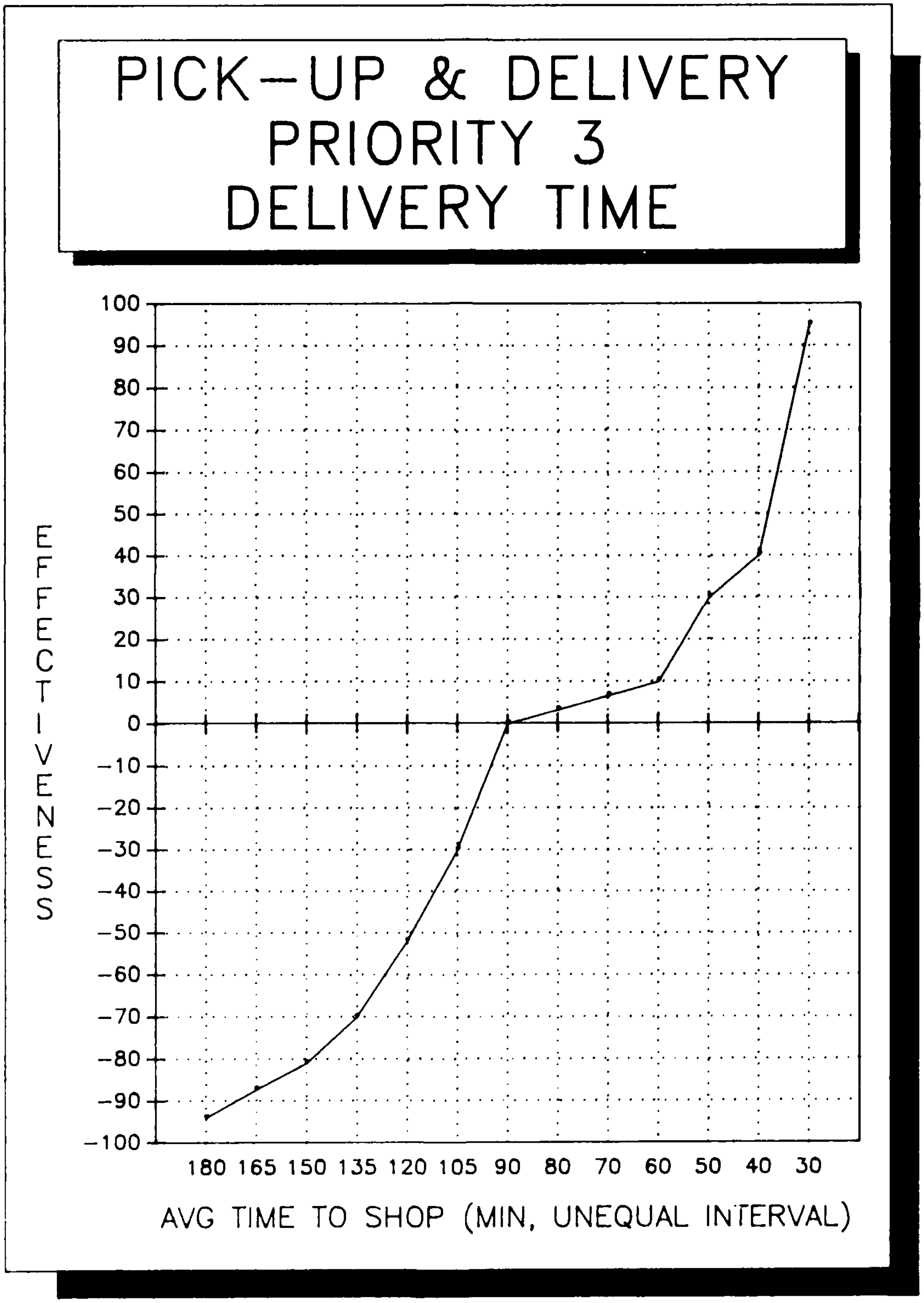


PICK-UP & DELIVERY
PICKUP OF TURN-INS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
35
30
25
20
15
10
5
0
NUMBER OF DELINQUENT TURN-INS

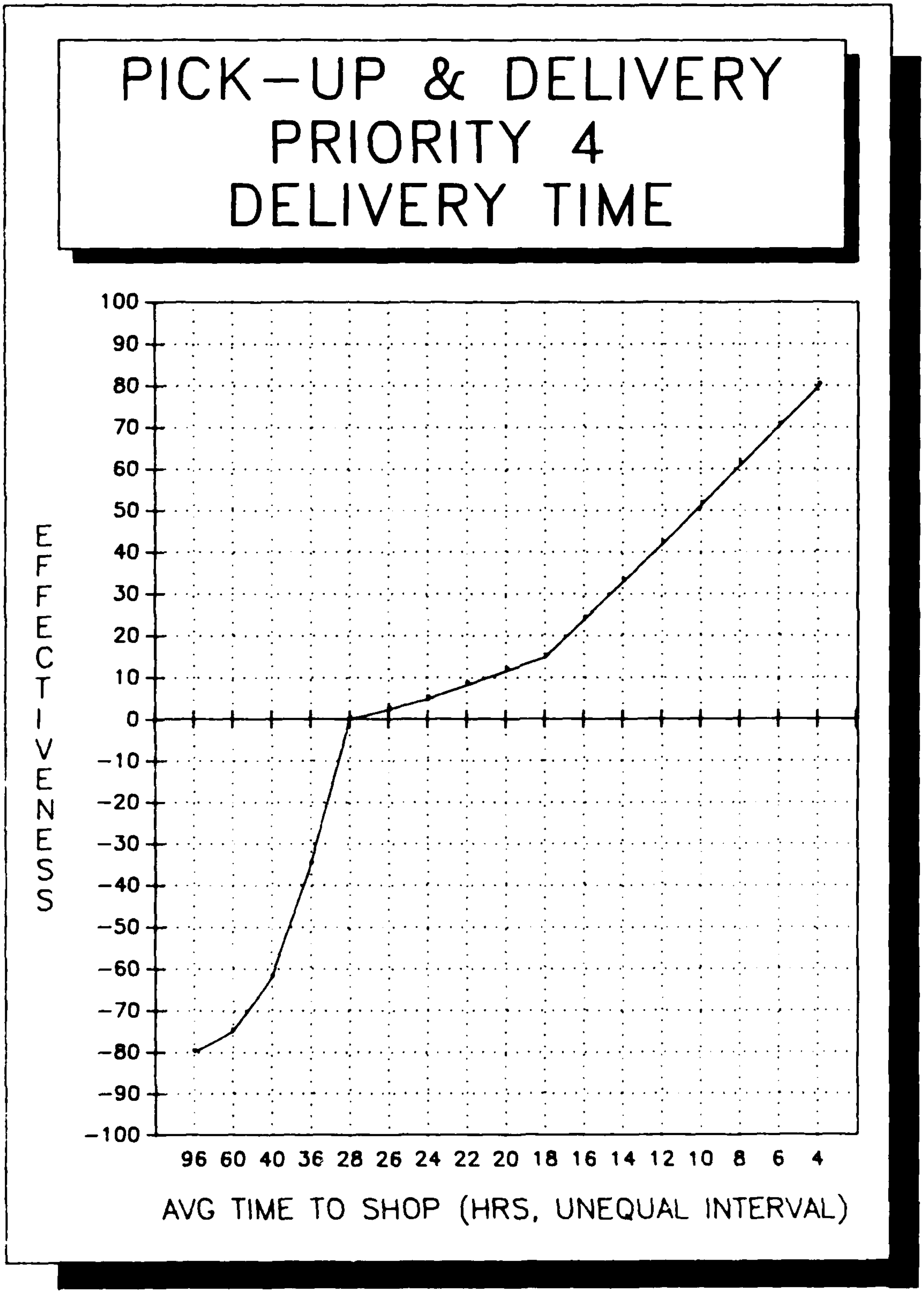


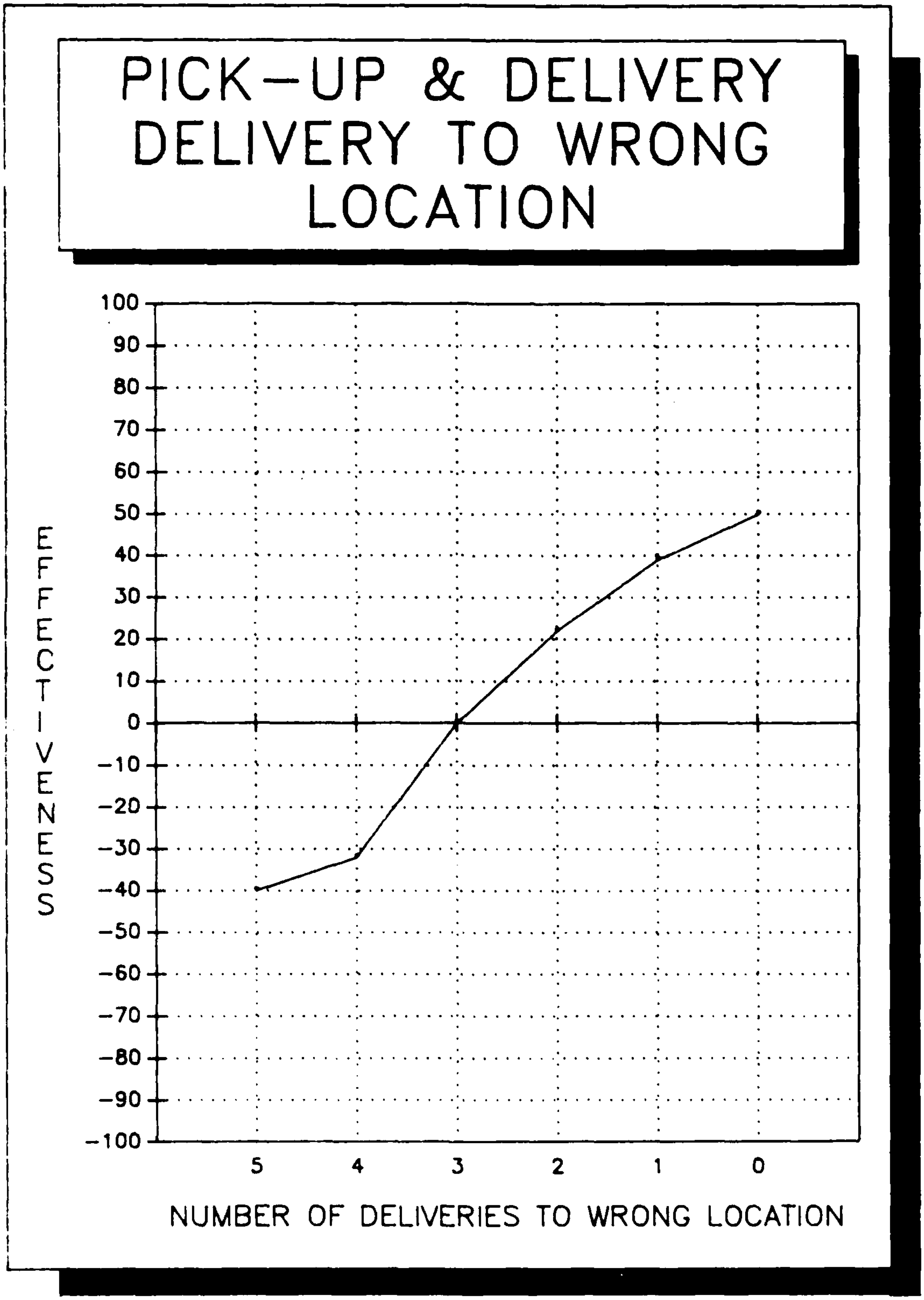
PICK-UP & DELIVERY
PRIORITY 2 (ISSUE)
DELIVERY TIME
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
120
105
90
75
60
45
30
15
AVG TIME TO SHOP (MIN,)

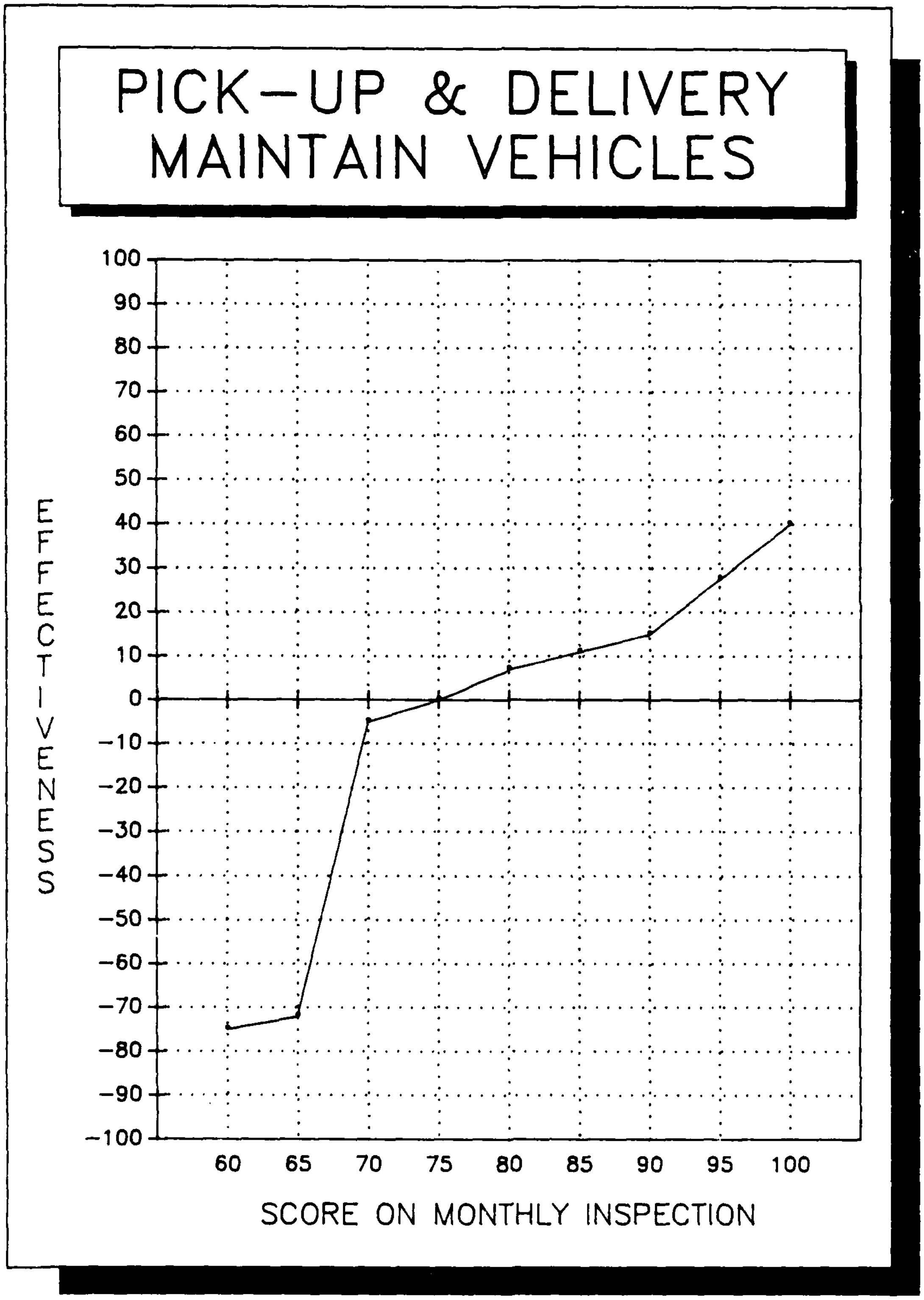

PICK-UP & DELIVERY
PRIORITY 2 (DOR)
DELIVERY TIME
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
180
110
90
80
70
60
50
40
30
AVG TIME TO SHOP (MIN, UNEQUAL INTERVALS)

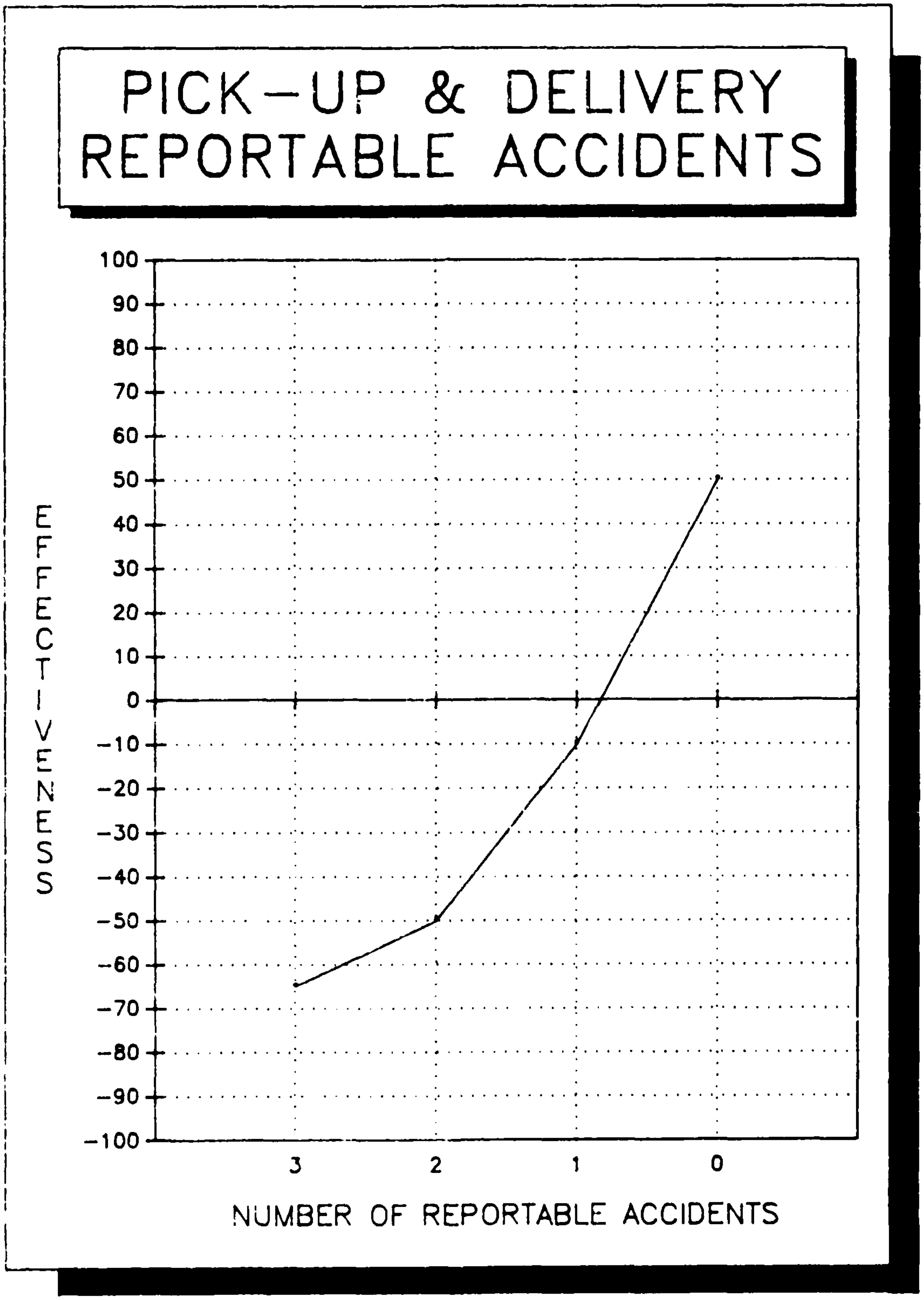


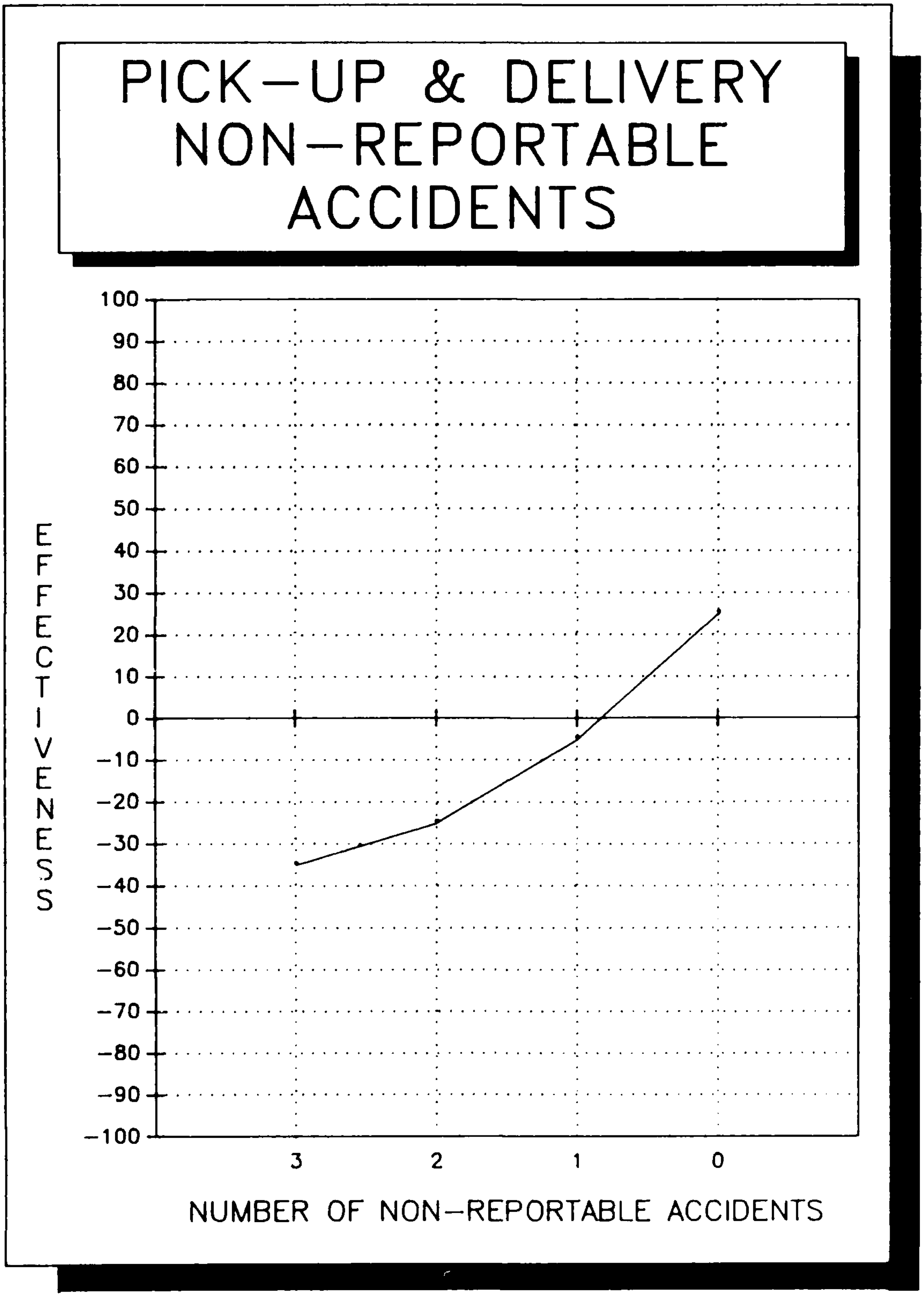
PICK-UP & DELIVERY
PRIORITY 3
DELIVERY TIME
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
EFFECTIVENESS
180 165 150 135 120 105 90 80 70 60 50 40 30
AVG TIME TO SHOP (MIN, UNEQUAL INTERVAL)

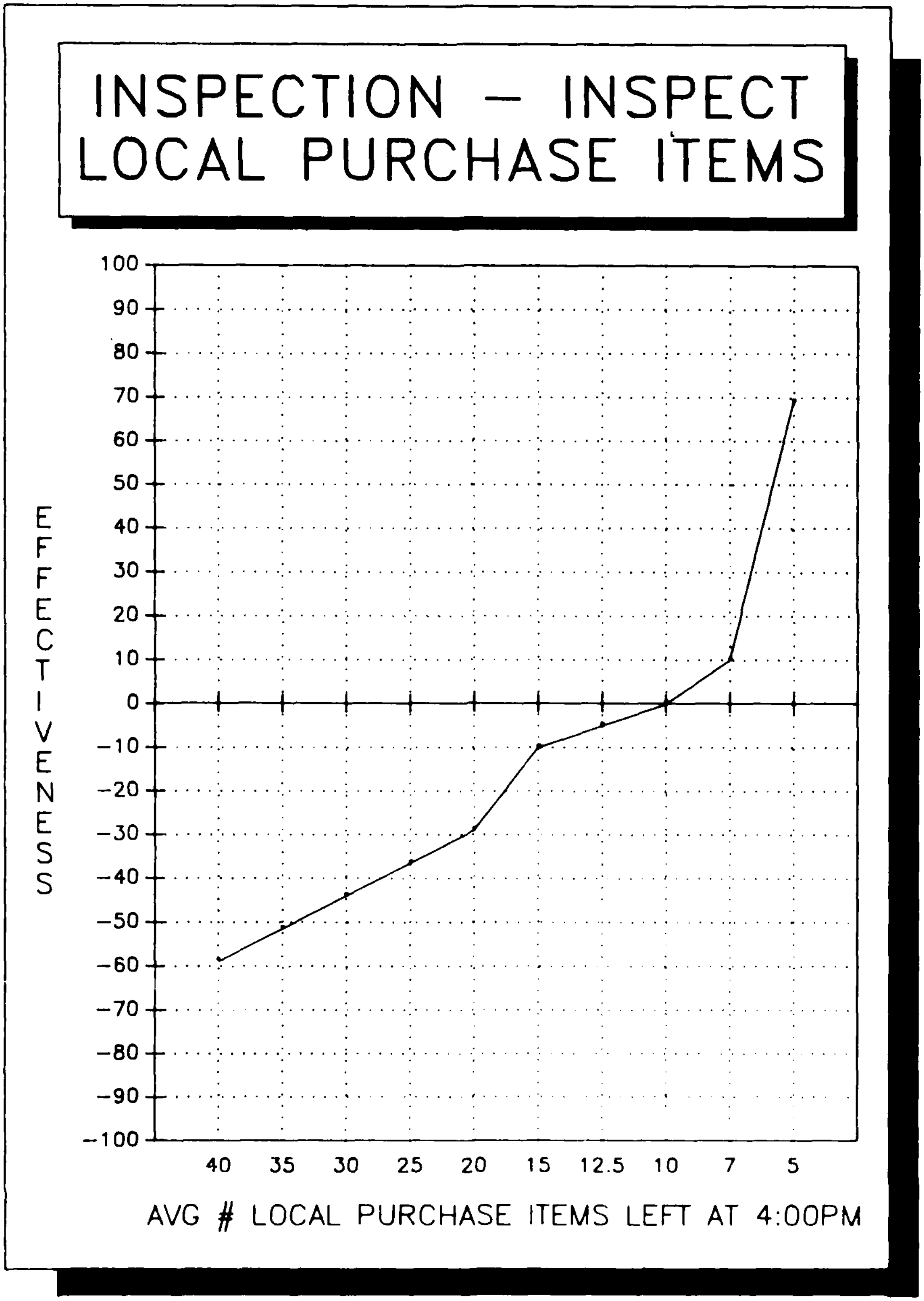

PICK-UP & DELIVERY
PRIORITY 4
DELIVERY TIME
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
96 60 40 36 28 26 24 22 20 18 16 14 12 10 8 6 4
AVG TIME TO SHOP (HRS, UNEQUAL INTERVAL)


PICK-UP & DELIVERY
DELIVERY TO WRONG LOCATION
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
5
4
3
2
1
0
NUMBER OF DELIVERIES TO WRONG LOCATION

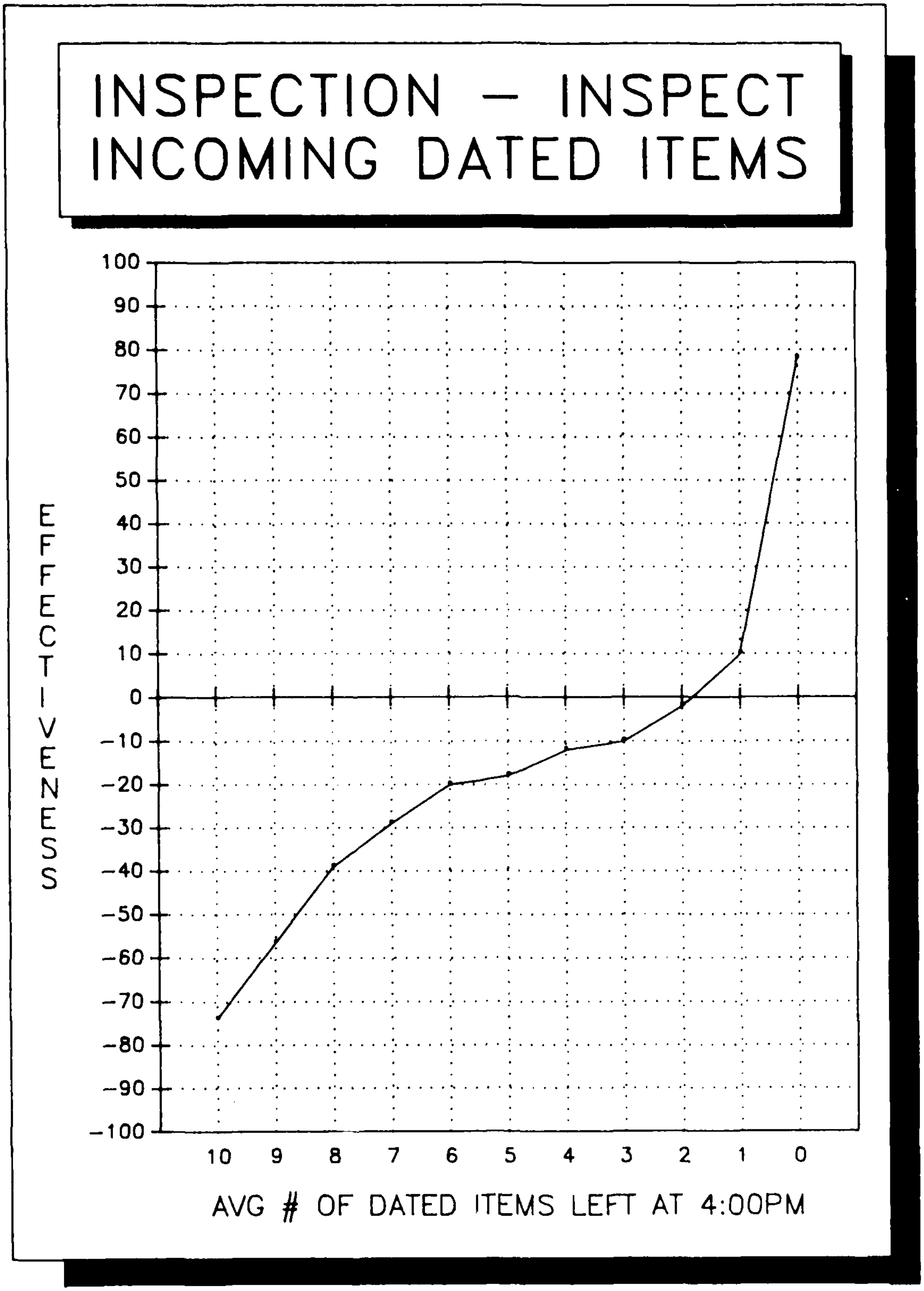

PICK-UP & DELIVERY
MAINTAIN VEHICLES
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
60
65
70
75
80
85
90
95
100
SCORE ON MONTHLY INSPECTION

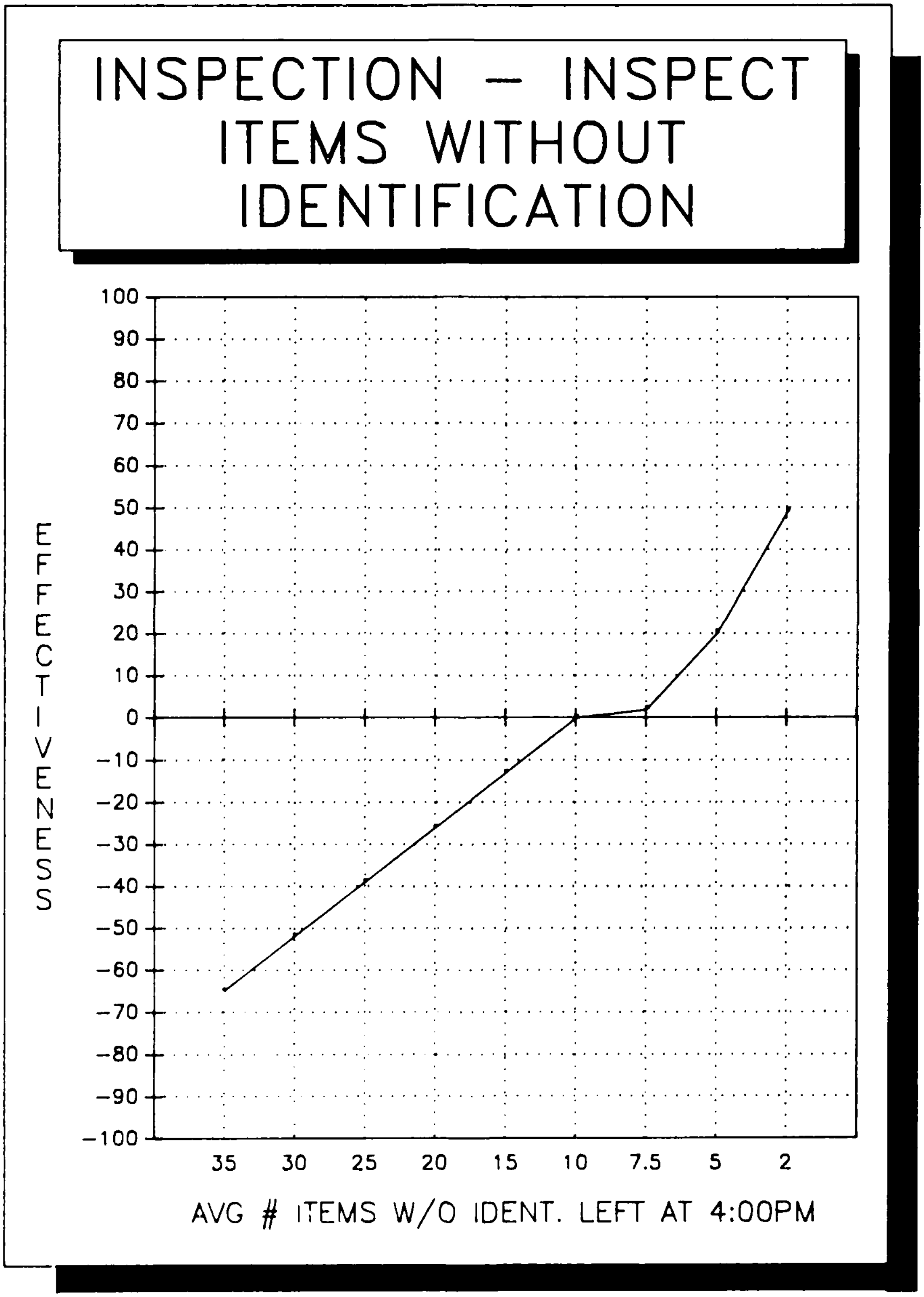

PICK-UP & DELIVERY
REPORTABLE ACCIDENTS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
3
2
1
0
NUMBER OF REPORTABLE ACCIDENTS

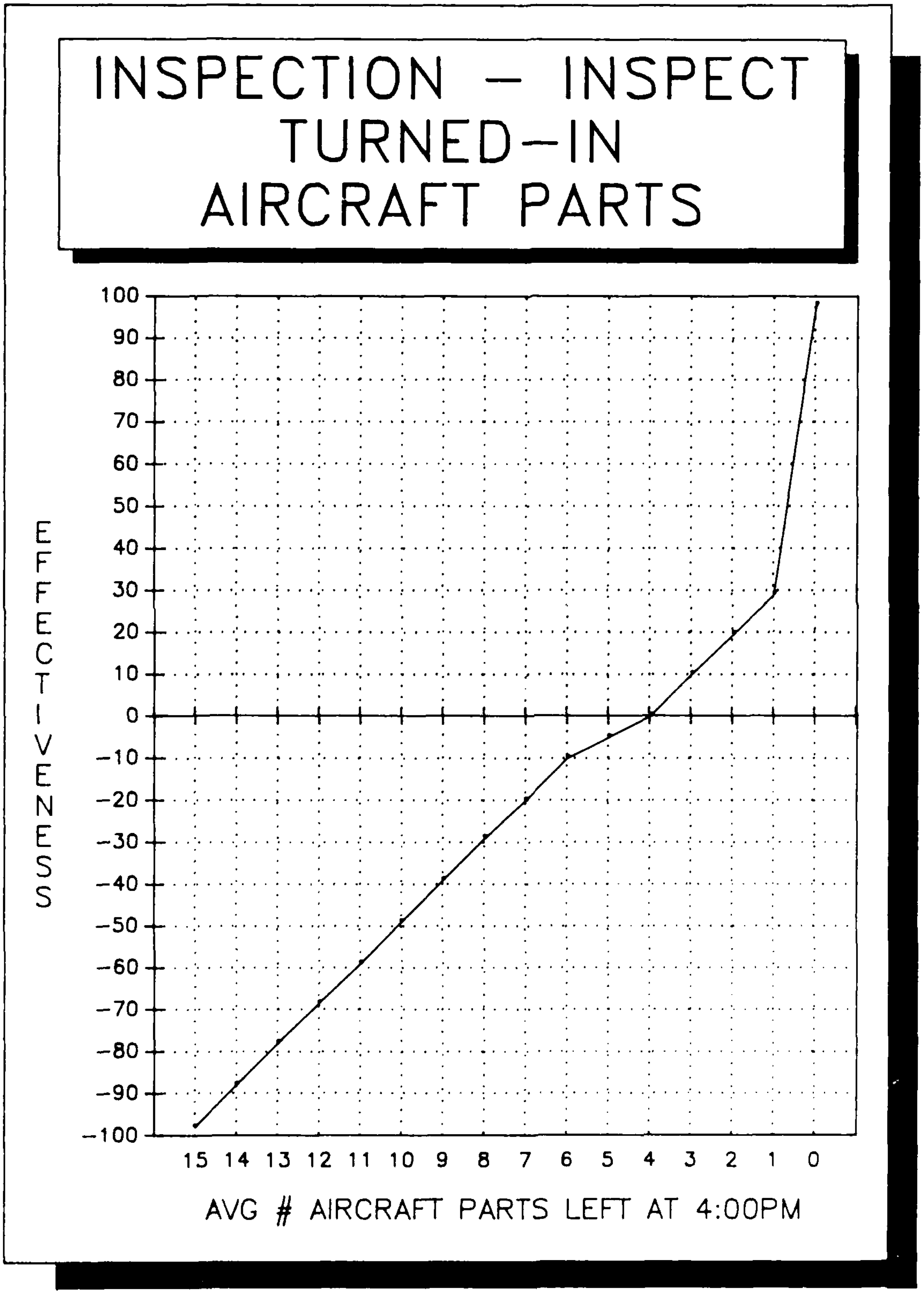

PICK-UP & DELIVERY
NON-REPORTABLE
ACCIDENTS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
3
2
1
0
NUMBER OF NON-REPORTABLE ACCIDENTS

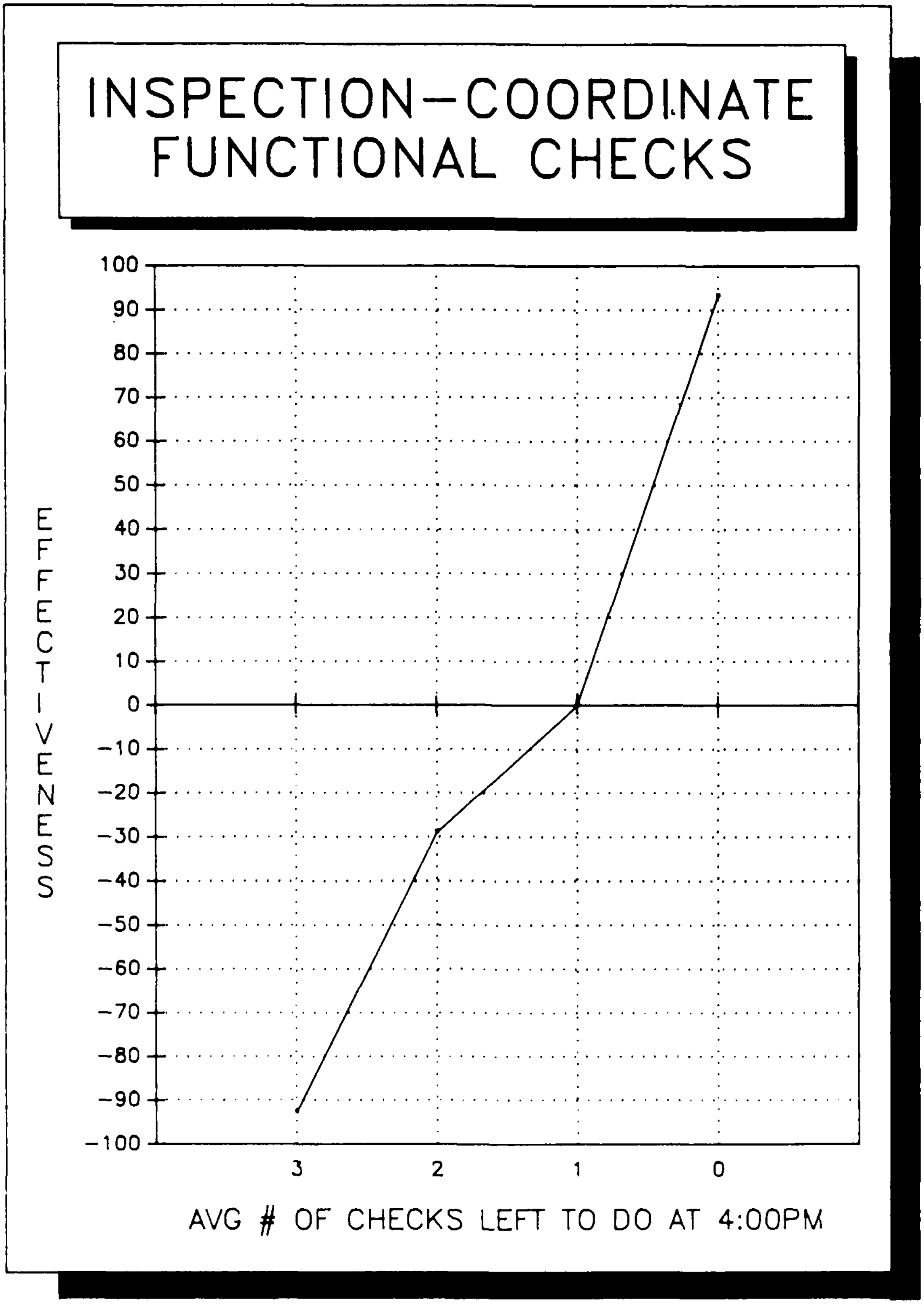


INSPECTION - INSPECT LOCAL PURCHASE ITEMS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
40
35
30
25
20
15
12.5
10
7
5
AVG # LOCAL PURCHASE ITEMS LEFT AT 4:00PM

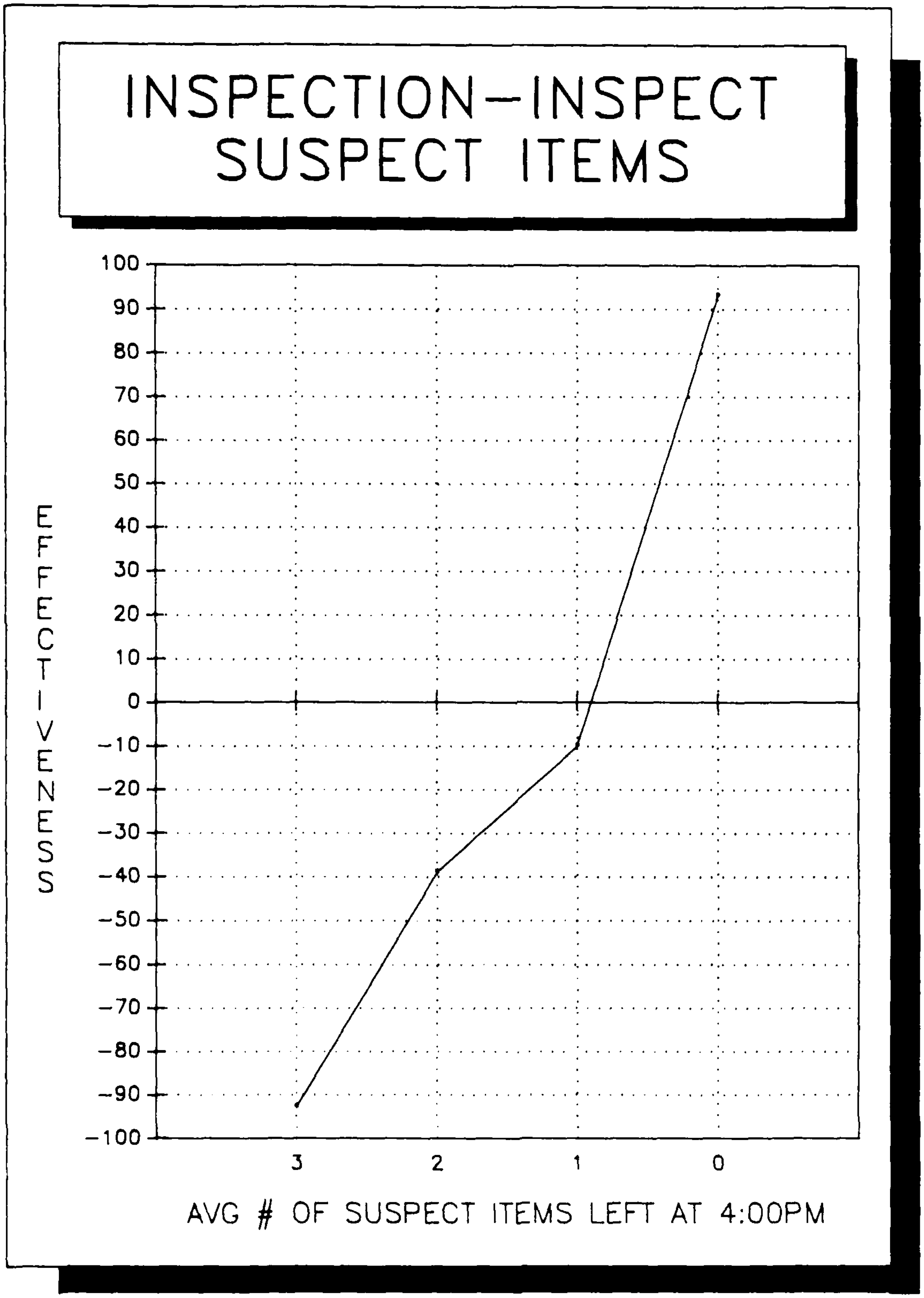

INSPECTION — INSPECT INCOMING DATED ITEMS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
10 9 8 7 6 5 4 3 2 1 0
AVG # OF DATED ITEMS LEFT AT 4:00PM

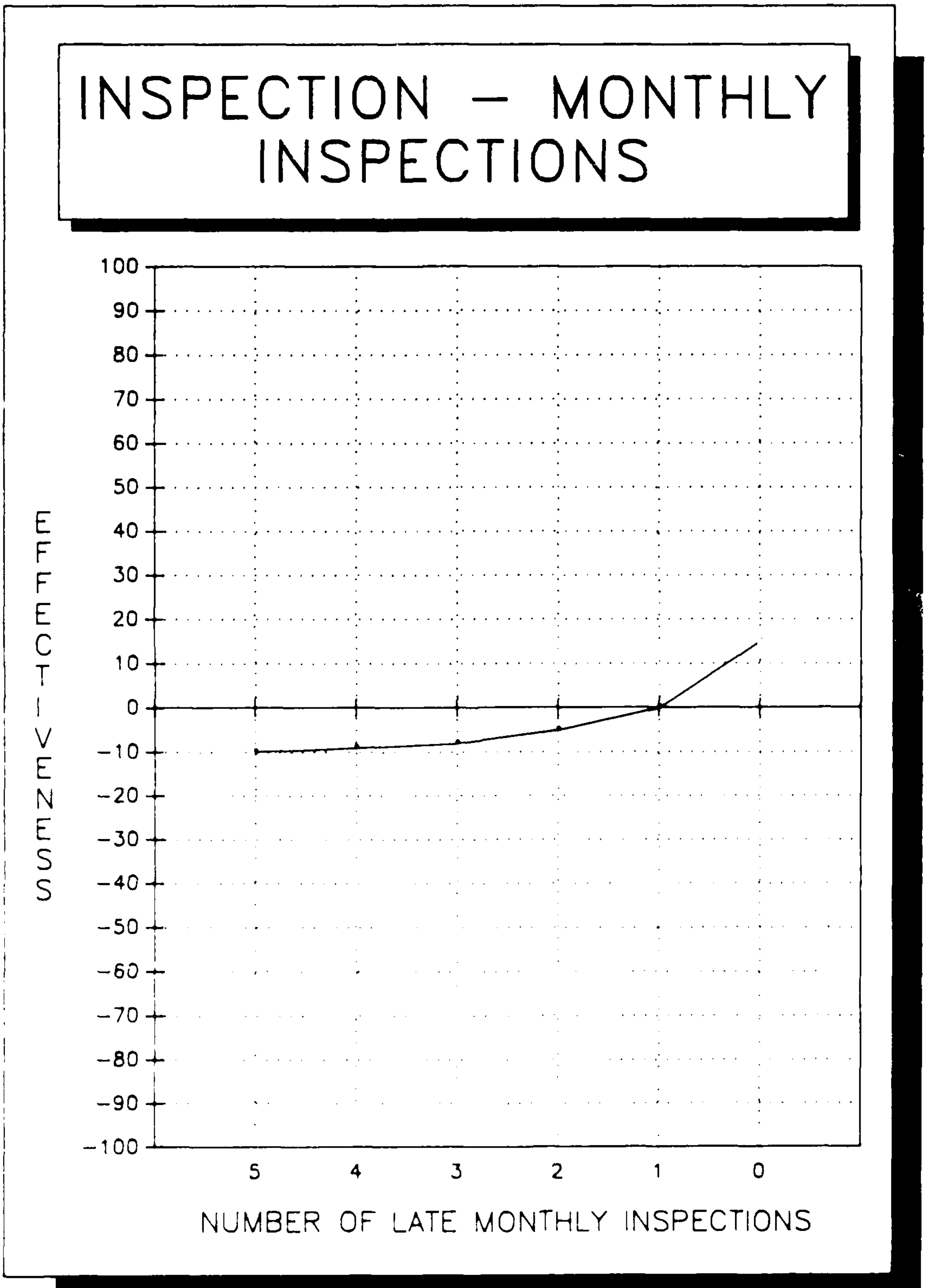

INSPECTION – INSPECT ITEMS WITHOUT IDENTIFICATION
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−100
35
30
25
20
15
10
7.5
5
2
AVG # ITEMS W/O IDENT. LEFT AT 4:00PM


INSPECTION – INSPECT TURNED–IN AIRCRAFT PARTS
100
90
80
70
60
50
40
30
20
10
0
–10
–20
–30
–40
–50
–60
–70
–80
–90
–100
EFFECTIVENESS
15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0
AVG # AIRCRAFT PARTS LEFT AT 4:00PM

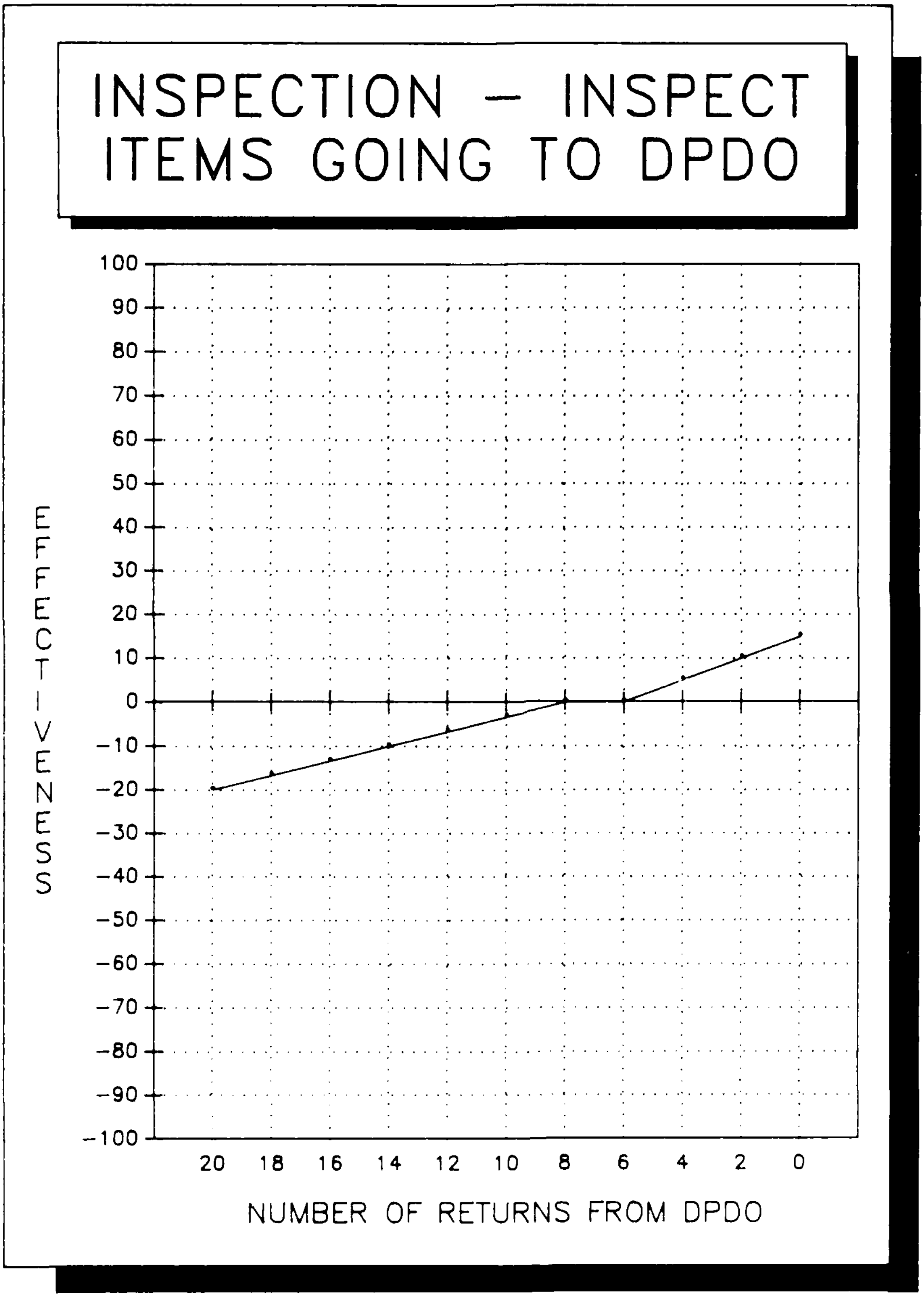


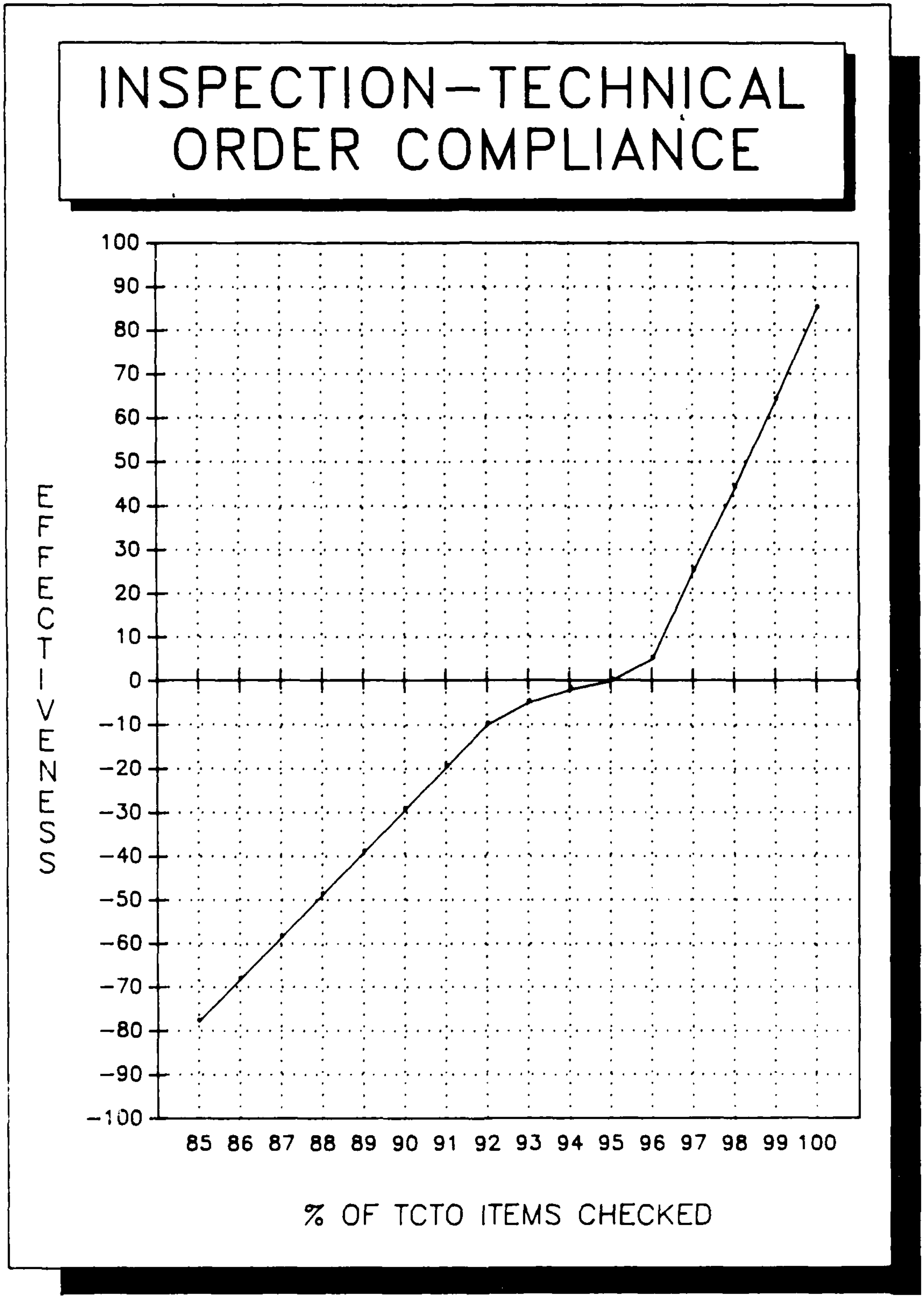
INSPECTION-COORDINATE
FUNCTIONAL CHECKS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
3
2
1
0
AVG # OF CHECKS LEFT TO DO AT 4:00PM

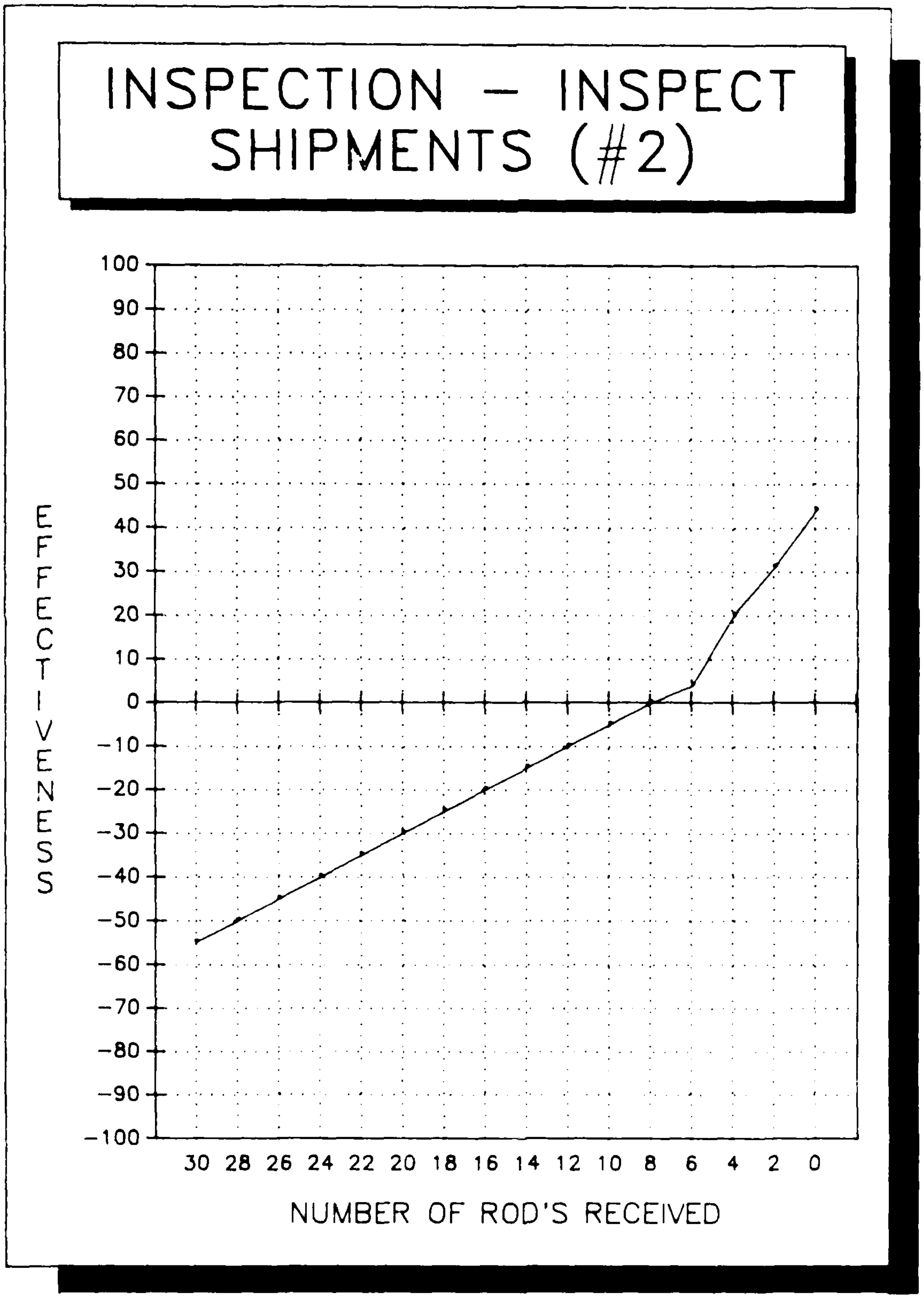

INSPECTION–INSPECT SUSPECT ITEMS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
–10
–20
–30
–40
–50
–60
–70
–80
–90
–100
3
2
1
0
AVG # OF SUSPECT ITEMS LEFT AT 4:00PM



INSPECTION – MONTHLY INSPECTIONS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
EFFECTIVENESS
5
4
3
2
1
0
NUMBER OF LATE MONTHLY INSPECTIONS

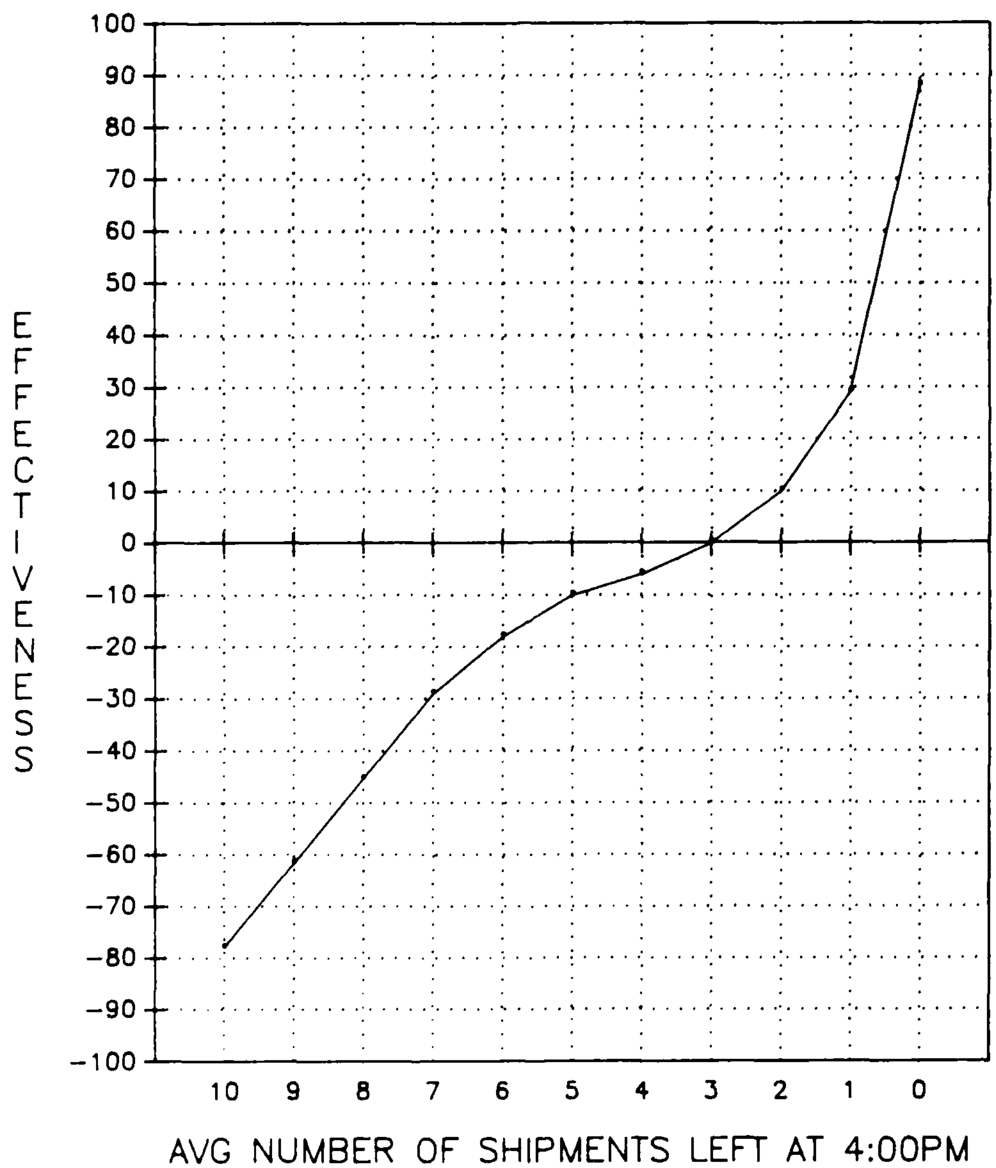

INSPECTION - INSPECT ITEMS GOING TO DPDO
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
20 18 16 14 12 10 8 6 4 2 0
NUMBER OF RETURNS FROM DPDO

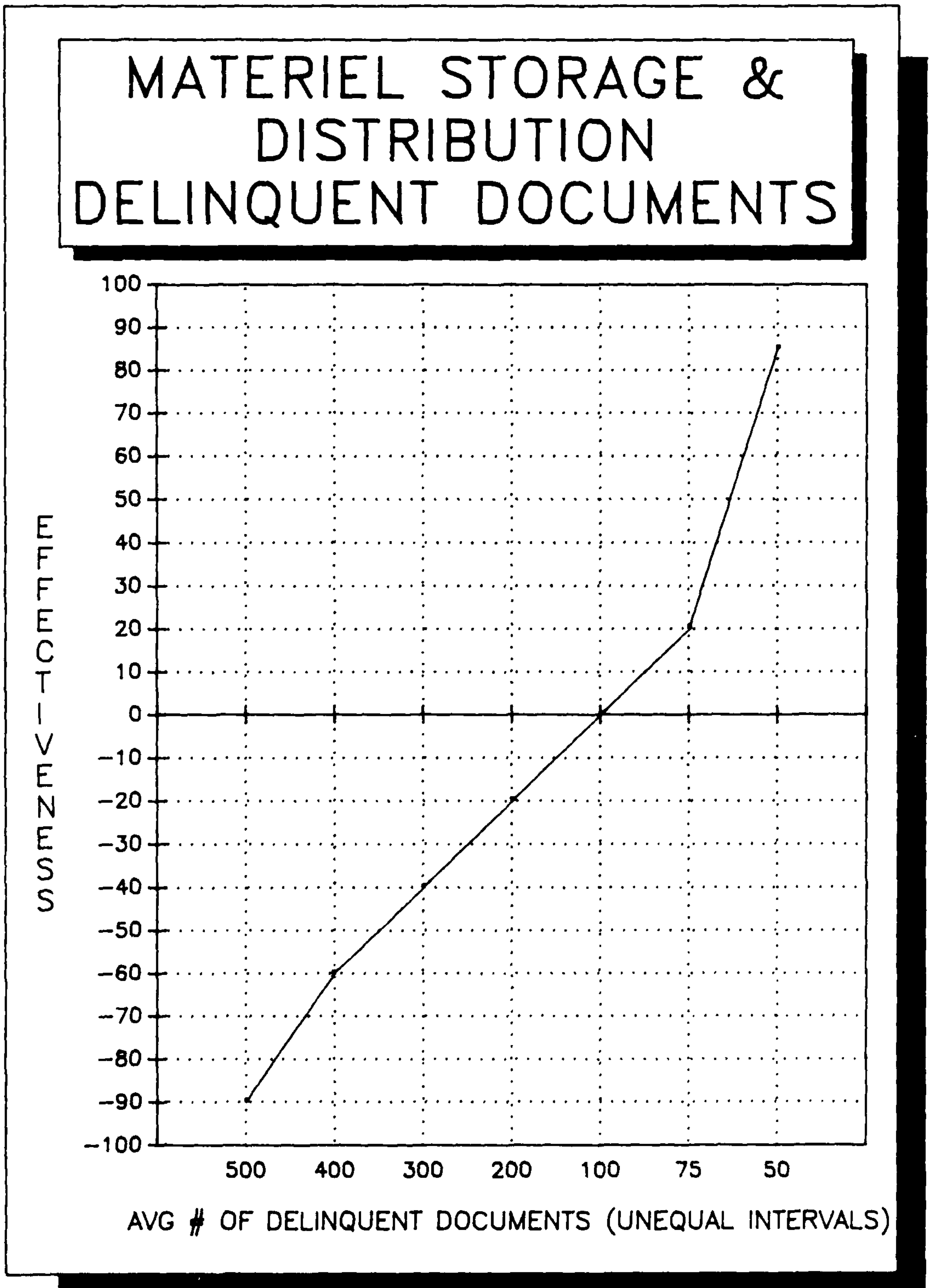

INSPECTION-TECHNICAL ORDER COMPLIANCE
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
% OF TCTO ITEMS CHECKED



INSPECTION - INSPECT SHIPMENTS (#2)
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
EFFECTIVENESS
30 28 26 24 22 20 18 16 14 12 10 8 6 4 2 0
NUMBER OF ROD'S RECEIVED

INSPECTION – INSPECT SHIPMENTS




MATERIEL STORAGE &
DISTRIBUTION
DELINQUENT DOCUMENTS
EFFECTIVENESS
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
500
400
300
200
100
75
50
AVG # OF DELINQUENT DOCUMENTS (UNEQUAL INTERVALS)

APPENDIX B: MS&D FEEDBACK REPORTS

APPENDIX B. MS&D FEEDBACK REPORTS

PRODUCTIVITY REPORT

MS&D BRANCH

INDICATOR AND EFFECTIVENESS DATA FOR MAY, 1986

BRANCH SUMMARY DATA

	INDICATOR DATA	EFFECTIVENESS SCORE	PERCENT OF MAX POSSIBLE
RECEIVING TOTAL		423	94%
STORAGE & ISSUE TOTAL		531	91%
PICKUP & DELIVERY TOTAL		387	65%
INSPECTION TOTAL		631	87%
DELINQUENT DOCUMENTS	58	65	76%
BRANCH GRAND TOTAL		2037	83%

APPENDIX B. (Cont.)

PRODUCTIVITY REPORT

MAY, 1986

RECEIVING SECTION

INDICATOR AND EFFECTIVENESS DATA

	INDICATOR DATA	EFFECTIVE-NESS SCORE	PERCENT OF MAX POSSIBLE
RECEIVE MATERIAL			
% IN-CHECKING ERRORS	.00	99	
DISTRIBUTE MATERIAL			
PRIOR. 2 DELIVERY TIME (MIN)	15.60	78	
PRIOR. 4 DELIVERY TIME (HRS)	2.57	87	
REFUSALS IN RECEIVING	0.00	75	
MONITOR REJECTS			
CLEARED DELINQUENT REJECTS	0.25	84	
TOTAL EFFECTIVENESS		423	94%

BRANCH LEVEL INDICATOR	INDICATOR DATA	EFFECTIVE-NESS SCORE
DELINQUENT DOCUMENTS	58	65

EFFECTIVENESS CHANGE FROM LAST MONTH

	LAST MONTH	THIS MONTH	CHANGE
RECEIVE MATERIAL			
% IN-CHECKING ERRORS	99	99	0
DISTRIBUTE MATERIAL			
PRIOR. 2 DELIVERY TIME (MIN)	80	78	-2
PRIOR. 4 DELIVERY TIME (HRS)	85	87	2
WHSE REFUSALS IN RECEIVING	75	75	0
MONITOR REJECTS			
CLEARED DELINQUENT REJECTS	85	84	-1

APPENDIX B. (Cont.)

RECEIVING SECTION

POTENTIAL EFFECTIVENESS GAIN

	INDICATOR DATA FROM	INDICATOR DATA TO	EFFECTIVE-NESS GAIN
RECEIVE MATERIAL			
% IN-CHECKING ERRORS	0	0	0
DISTRIBUTE MATERIAL			
PRIOR. 2 DELIVERY TIME (MIN)	16	0	22
PRIOR. 4 DELIVERY TIME (HRS)	3	0	3
WHSE REFUSALS IN RECEIVING	0	0	0
MONITOR REJECTS			
CLEARED DELINQUENT REJECTS	1	0	1

APPENDIX B. (Cont.)

PRODUCTIVITY REPORT: MAY, 1986

STORAGE AND ISSUE SECTION

INDICATOR AND EFFECTIVENESS DATA

	INDICATOR DATA	EFFECTIVENESS SCORE	PERCENT OF MAX POSSIBLE
STORE ITEMS IN WAREHOUSE			
PROPER LOCATION: # REFUSALS	0	85	
PROPER PROCEDURE: # FINDINGS	18	4	
WAREHOUSE MAINTENANCE			
MAINTAIN LOCATION: % OFF R36	98.13	83	
ISSUE REQUESTED ITEMS			
PRIOR. 2 DELIVERY TIME (MIN)	9.96	100	
PRIOR. 3 DELIVERY TIME (MIN)	17.40	99	
PRIOR. 4 DELIVERY TIME (HRS)	1.39	90	
RESPOND TO SURVEILLANCES			
# REPEAT FINDINGS	0	70	
TOTAL EFFECTIVENESS		531	91%

BRANCH LEVEL INDICATOR	INDICATOR DATA	EFFECTIVENESS SCORE
DELINQUENT DOCUMENTS	58	65

EFFECTIVENESS CHANGE FROM LAST MONTH

	LAST MONTH	THIS MONTH	CHANGE
PROPER LOCATION: # REFUSALS	85	85	0
PROPER PROCEDURE: # FINDINGS	34	4	-30
MAINTAIN LOCATION: % OFF R36	87	83	-4
PRIOR. 2 DELIVERY TIME (MIN)	100	100	0
PRIOR. 3 DELIVERY TIME (MIN)	98	99	1
PRIOR. 4 DELIVERY TIME (HRS)	88	90	2
# REPEAT FINDINGS	70	70	0

APPENDIX B. (Cont.)

STORAGE AND ISSUE SECTION

POTENTIAL EFFECTIVENESS GAIN

	INDICATOR FROM	INDICATOR TO	EFFECTIVENESS GAIN
STORE ITEMS IN WAREHOUSE			
PROPER LOCATION: #REFUSALS	0	0	0
PROPER PROCEDURE: # FINDINGS	18	14	8
WAREHOUSE MAINTENANCE			
MAINTAIN LOCATION: % OFF R36	98	100	9
ISSUE REQUESTED ITEMS			
PRIOR. 2 DELIVERY TIME (MIN)	10	0	0
PRIOR. 3 DELIVERY TIME (MIN)	17	4	0
PRIOR. 4 DELIVERY TIME (HRS)	1	0	0
RESPOND TO SURVEILLANCES			
# REPEAT FINDINGS	0	0	0

PRODUCTIVITY REPORT: MAY, 1986

PICKUP & DELIVERY SECTION

	INDICATOR DATA	EFFECTIVENESS SCORE	PERCENT OF MAX POSSIBLE
PICK UP TURN-INS			
# DELINQUENT TURN-INS	0	65	
DELIVER ITEMS PROPERLY			
PRIOR. 2 I DELIVERY TIME (MIN)	27.60	41	
PRIOR. 2 D DELIVERY TIME (MIN)	34.20	68	
PRIOR. 3 DELIVERY TIME (MIN)	39.00	46	
PRIOR. 4 DELIVERY TIME (HRS)	15.34	27	
# DELIVERED WRONG LOCATION	0	50	
MAINTAIN VEHICLES			
VEHICLE INSPECTION SCORE	90.00	15	
# REPORTABLE ACCIDENTS	0	50	
# NONREPORTABLE ACCIDENTS	0	25	
TOTAL EFFECTIVENESS FOR MONTH		387	65%

BRANCH LEVEL INDICATOR	INDICATOR DATA	EFFECTIVENESS SCORE
DELINQUENT DOCUMENTS	58	65

EFFECTIVENESS CHANGE FROM LAST MONTH

	LAST MONTH	THIS MONTH	CHANGE
# DELINQUENT TURN-INS	65	65	0
PRIOR. 2 I DELIVERY TIME (MIN)	66	41	-25
PRIOR. 2 D ELIVERY TIME (MIN)	40	68	28
PRIOR. 3 DELIVERY TIME (MIN)	55	46	-9
PRIOR. 4 DELIVERY TIME (HRS)	27	27	0
# DELIVERED WRONG LOCATION	50	50	0
VEHICLE INSPECTION SCORE	30	15	-15
# REPORTABLE ACCIDENTS	50	50	0
# NONREPORTABLE ACCIDENTS	25	25	0

APPENDIX B. (Cont.)

PICKUP & DELIVERY SECTION

POTENTIAL EFFECTIVENESS GAIN

	INDICATOR FROM	INDICATOR TO	EFFECTIVE-NESS GAIN
PICK UP TURN-INS			
# DELINQUENT TURN-INS	0	0	0
DELIVER ITEMS PROPERLY			
PRIOR. 2 I DELIVERY TIME (MIN)	28	8	59
PRIOR. 2 D DELIVERY TIME (MIN)	34	4	27
PRIOR. 3 DELIVERY TIME (MIN)	39	9	49
PRIOR. 4 DELIVERY TIME (HRS)	15	0	53
# DELIVERED WRONG LOCATION	0	0	0
MAINTAIN VEHICLES			
VEHICLE INSPECTION SCORE	90	98	20
# REPORTABLE ACCIDENTS	0	0	0
# NONREPORTABLE ACCIDENTS	0	0	0

APPENDIX B. (Cont.)

PRODUCTIVITY REPORT: MAY, 1986

INSPECTION SECTION

INDICATOR AND EFFECTIVENESS DATA

	INDICATOR DATA	EFFECTIVE-NESS SCORE	PERCENT OF MAX POSSIBLE
AVG # LOCAL PURCHASE ITEMS LEFT TO BE INSPECTED	2.84	69	
AVG # INCOMMING DATED ITEMS LEFT TO BE INSPECTED	0.00	78	
AVG # ITEMS W/O IDENTIFICATION LEFT TO BE INSPECTED	5.16	19	
AVG # TURNED-IN AIRCRAFT PARTS LEFT TO BE INSPECTED	0.11	91	
AVG # FUNCTIONAL CHECKS LEFT TO BE COORDINATED	0.11	83	
AVG # SUSPECT ITEMS LEFT TO BE INSPECTED	0.00	93	
# LATE MONTHLY INSPECTIONS	0.00	15	
INSPECT DPDO MATERIAL (# RET.)	0.00	15	
TECH. ORDERS: % TCTOs CHECKED	100	85	
SHIPMENTS: # RODS	6.00	4	
SHIPMENTS: AVG # SHIPMENTS LEFT TO BE INSPECTED	0.16	79	
TOTAL EFFECTIVENESS		631	87%

BRANCH LEVEL INDICATOR	INDICATOR DATA	EFFECTIVENESS SCORE
DELINQUENT DOCUMENTS	58	65

APPENDIX B. (Cont.)

INSPECTION SECTION

EFFECTIVENESS CHANGE FROM LAST MONTH

	LAST MONTH	THIS MONTH	CHANGE
AVG # LOCAL PURCHASE ITEMS LEFT TO BE INSPECTED	-6	69	75
AVG # INCOMMING DATED ITEMS LEFT TO BE INSPECTED	78	78	0
AVG # ITEMS W/O IDENTIFICATION LEFT TO BE INSPECTED	20	19	-1
AVG # TURNED-IN AIRCRAFT PARTS LEFT TO BE INSPECTED	98	91	-7
AVG # FUNCTIONAL CHECKS LEFT TO BE COORDINATED	-17	83	100
AVG # SUSPECT ITEMS LEFT TO BE INSPECTED	93	93	0
# LATE MONTHLY INSPECTIONS	15	15	0
INSPECT DPDO MATERIAL (# RET.)	15	15	0
TECH ORDERS: % TCTOs CHECKED	85	85	0
SHIPMENTS: # RODS	-10	4	14
SHIPMENTS: AVG # SHIPMENTS LEFT TO BE INSPECTED	83	79	-4

APPENDIX B. (Concluded)

INSPECTION SECTION

POTENTIAL EFFECTIVENESS GAIN

	INDICATOR FROM	INDICATOR TO	EFFECTIVENESS GAIN
AVG # LOCAL PURCHASE ITEMS LEFT TO BE INSPECTED	3	0	0
AVG # INCOMMING DATED ITEMS LEFT TO BE INSPECTED	0	0	0
AVG # ITEMS W/O IDENTIFICATION LEFT TO BE INSPECTED	5	0	30
AVG # TURNED-IN AIRCRAFT PARTS LEFT TO BE INSPECTED	0	0	7
AVG # FUNCTIONAL CHECKS LEFT TO BE COORDINATED	0	0	10
AVG # SUSPECT ITEMS LEFT TO BE INSPECTED	0	0	0
# LATE MONTHLY INSPECTIONS	0	0	0
INSPECT DPDO MATERIAL (# RET.)	0	0	0
TECH. ORDERS: % TCTOs CHECKED	100	100	0
SHIPMENTS: # RODS	6	0	40
SHIPMENTS: AVG # SHIPMENTS LEFT TO BE INSPECTED	0	0	9

END
9-87
DTIC